

WO0032800

Publication Title:

ARTIFICIAL MATRIX ATTACHMENT REGION FOR INCREASING
EXPRESSION OF GENES INTRODUCED IN PLANT CELLS

Abstract:

1369 Abstract of WO0032800

Synthetic DNA molecule is useful as matrix attachment region to increase expression of genes introduced in transformed plants.

Data supplied from the esp@cenet database - Worldwide

Courtesy of <http://v3.espacenet.com>

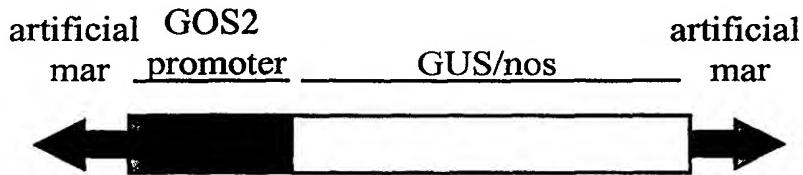


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/82, 15/29, 15/67, 15/63, 5/10		A1	(11) International Publication Number: WO 00/32800 (43) International Publication Date: 8 June 2000 (08.06.00)
(21) International Application Number: PCT/US99/28123		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 30 November 1999 (30.11.99)			
(30) Priority Data: 60/110,437 1 December 1998 (01.12.98) US			
(71) Applicant: DOW AGROSCIENCES LLC [US/US]; 9330 Zionsville Road, Indianapolis, IN 46268 (US).			
(72) Inventors: VAN DER GEEST, Apolonia, H.M.; Roghorst 365, NL-6708 KX Wageningen (NL). AINLEY, W., Michael; 1474 Clearwater Court, Carmel, IN 46032 (US). COWEN, Neil, M.; 990 Tillson Drive, Zionsville, IN 46077 (US). WELTER, Mary, E.; 5333 Guilford Avenue, Indianapolis, IN 46220 (US). WOOSLEY, Aaron, T.; 8906 Tanner Drive, Fishers, IN 46038 (US).			
(74) Agent: STUART, Donald, R.; Dow AgroSciences LLC, 9330 Zionsville Road, Indianapolis, IN 46268 (US).			

(54) Title: ARTIFICIAL MATRIX ATTACHMENT REGION FOR INCREASING EXPRESSION OF GENES INTRODUCED IN PLANT CELLS

Rice Construct



(57) Abstract

Synthetic DNA molecule is useful as matrix attachment region to increase expression of genes introduced in transformed plants.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**ARTIFICIAL MATRIX ATTACHMENT REGION FOR INCREASING
EXPRESSION OF GENES INTRODUCED IN PLANT CELLS**

The present invention relates to plant molecular
5 biology, and in particular to technology for enhancing
the expression of genes introduced in transformed plant
cells.

Through the use of recombinant DNA technology and
genetic engineering, it has become possible to introduce
10 desired DNA sequences into plant cells to allow for the
expression of proteins of interest. Plants with
genetically engineered traits, such as, for example,
insect resistance, disease resistance, drought
resistance, herbicide resistance, or metabolic
15 alterations that increase or modify production of useful
plant products, offer great promise of improving
agriculture.

Obtaining desired levels of expression of DNA
introduced into plant cells remains a challenge. One
20 problem, referred to as "position effect" variation, is
the variation in expression of the same gene in
independent transformants. The use of naturally
occurring DNA sequences called matrix attachment regions
or scaffold attachment regions to combat this problem was
25 proposed in U.S. Patent 5,773,689 and in WO 94/24293.

The present invention provides a novel synthetic DNA
molecule comprising bp 11 to 309 of SEQ ID NO: 1 that is
useful as a matrix attachment region to increase
expression of genes introduced in transformed plants.

In another of its aspects, the invention provides a
30 DNA construct comprising, in the 5' to 3' direction: a
transcription initiation region functional in plant
cells, a structural gene operatively associated with the
transcription initiation region, a 3' untranslated
region, and a matrix attachment region comprised of bp 11
35 to 309 of SEQ ID NO: 1 positioned either 5' to said
transcription initiation region or 3' to said structural

gene. In a preferred embodiment, a first matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 is positioned 5' to said transcription initiation region and a second matrix attachment region comprised of 5 bp 11 to 309 of SEQ ID NO: 1 is positioned 3' to said 3' untranslated region.

In a particularly preferred embodiment, the matrix attachment region of the invention comprises two or more tandem copies of bp 11 to 309 of SEQ ID NO:1.

10

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram showing the strategy for assembling the artificial MAR of SEQ ID NO:1.

15

FIG. 2 is a schematic representation of the rice transformation construct ArGOS2Af, which contains the MAR dimer.

FIG. 3 is a graph comparing relative GUS activity for multiple rice transformation events using non-MAR containing construct GOS2 and construct ArGOS2Af, which contains an artificial MAR dimer.

20

FIG. 4 is a graph showing the effect of the artificial MAR on ranges of expression of the GUS reporter gene in transgenic rice plants.

25

FIG. 5 is a schematic representation of the *Arabidopsis* transformation constructs ArAct2Af and aaaaAfAct2Af. ArAct2Af contains copies of the MAR dimer in opposite orientations flanking the reporter gene. AfAct2Af contains copies of the MAR dimer in the same orientation flanking the reporter gene.

30

FIG. 6 is a graph comparing relative GUS activity for multiple *Arabidopsis* transformation events using non-MAR containing construct Act2 and constructs ArAct2Af and AfAct2Af, which contain an artificial MAR dimer.

35

FIG. 7 is a graph showing the effect of the artificial MAR on the range of expression of the GUS reporter gene in transgenic *Arabidopsis* plants.

DESCRIPTION OF THE SEQUENCES

SEQ ID NO:1 describes the artificial MAR of the invention.

SEQ ID NOS:2 to 4 describe ARBP sites.

SEQ ID NO:5 describes an ATF site.

5 SEQ ID NO:6 describes a BEAF-32 site.

SEQ ID NOS:7 to 9 describe topoisomerase II sites.

SEQ ID NO:10 describes an unwinding sequence.

SEQ ID NOS:11 to 17 describe SATB I sites.

SEQ ID NO:18 describes exemplary bending DNA.

10 SEQ ID NO:19 describes an exemplary A/T tract.

SEQ ID NO:20 describes synthetic MAR-A.

SEQ ID NO:21 describes synthetic MAR-B.

SEQ ID NO:22 describes synthetic MAR-C.

SEQ ID NO:23 describes synthetic MAR-D.

15 SEQ ID NO:24 describes synthetic MAR-E.

SEQ ID NO:25 describes synthetic MAR-F.

SEQ ID NO:26 describes the 3' MAR dimer in
pArGOS2Af-hpt and ArAct2Af-bin

SEQ ID NO:27 describes rice transformation vector
20 pGOS2-hpt.

SEQ ID NO:28 describes rice transformation vector
pArGOS2Af-hpt.

SEQ ID NO:29 describes the 5' MAR dimer in
pArGOS2Af-hpt and ArAct2Af-bin.

25 SEQ ID NO:30 describes dicot transformation vector
pAct2-bin.

SEQ ID NO:31 describes dicot transformation vector
pArAct2Af-bin.

30 SEQ ID NO:32 describes dicot transformation vector
pAfAct2Af-bin.

SEQ ID NO:33 describes the 3' MAR dimer in
pAfAct2Af-bin.

DETAILED DESCRIPTION OF THE INVENTION

Eukaryotic nuclei are highly organized structures in
35 which the entire genetic information has to be accessible
in an orderly manner for replication, transcription and

other cellular events (Lewin, 1994; Dillon and Grosveld, 1994; Jackson, 1995; Wolffe, 1994). Genes are typically organized in chromatin loops of various sizes that are attached to the proteinaceous nuclear matrix at locations known as matrix attachment regions (MARs). MARs are often located in non-transcribed regions of genes and are thought to form the physical boundaries of individual DNA loops. In several cases, MARs were shown to reduce position effect in transgenic organisms. The chicken lysozyme MAR was shown to increase expression, reduce variance and make expression of an adjacent gene copy number dependent in stably transfected cells (Stief et al., 1989) in transgenic mice (Bonifer et al., 1990, McKnight et al., 1992) and in transgenic tobacco plants (Mlynárová et al., 1994; Mlynárová et al., 1995).

However, not all MARs have these effects on gene expression. Two minimal *Drosophila* MARs (one located between the histone H1 and H3 genes, and the other near the heat shock HSP70 genes) stimulated expression more than 10-fold in stably transformed cells, but the presence of these MARs did not reduce position effect (Poljak et al., 1994). MARs from the apolipoprotein domain increased expression and reduced position effect in low-copy transformants, but expression in multicopy transformants was strongly repressed. (Kalos and Fournier, 1995). When a *Drosophila ftz* MAR was placed in a different chromosomal location, it did not reorganize chromatin structure and the chromatin fragment containing the MAR could be easily eluted from the nucleus, indicating that introduced MARs do not necessarily form chromatin domains (Eggert and Jack, 1991). In contrast, MARs flanking the immunoglobulin m heavy chain locus enhancer were required for high levels of expression and the formation of an extended DNase I sensitive domain in transgenic B lymphocytes, but not in stably transfected tissue culture cells (Forrester et al., 1994).

Results using MARs in transgenic plants have been similarly complex (Spiker and Thompson, 1996). A yeast MAR increased expression levels in stably transformed tobacco callus lines, but no correlation between copy 5 number and expression level could be found (Allen et al., 1993). In contrast, the MAR element from the soybean heat shock gene *Gmhsp17.6-L* was shown to be capable of increasing expression levels but had little effect on variability (Schöffl et al., 1993). A soybean MAR 10 flanking a reporter gene construct reduced variability of expression when compared to a construct lacking MARs, but also reduced expression levels when present 5' and 3' of a reporter gene construct in transgenic tobacco callus (Breyne et al., 1992). It is possible that individual 15 MARs can have different functional and structural properties in addition to their matrix binding ability (Breyne et al., 1994).

MARs are usually 300 to 2000 base pairs in length, are rich in adenine and thymine residues and often 20 contain certain conserved sequence elements and structural features. Most MARs described in the literature are not obtained from plants, but it has been well documented that MARs from other organisms bind plant scaffolds and vice versa (Dietz et al., 1994; Breyne et 25 al., 1992).

Table 1 describes sequence elements present in MARs described in the literature, including the following plant MARs: soybean heat shock protein gene MAR, (Schöffl et al., 1993); a petunia MAR (Dietz et al., 1994); the 30 pea plastocyanin gene MAR (Slatter et al., 1991); the maize *Adh1* gene 5' and 3' MARs (Avramova and Bennetzen, 1993; Avramova et al., 1995); the *b*-phaseolin gene 5' and 3' MARs (van der Geest et al., 1994).

Table 1

element	characteristics	sequence	SEQ ID NO:	reported in plant MARs
ARBP	ARBP (attachment-region binding protein) from chicken which binds MARs from <i>Drosophila</i> , mouse, chicken and human genes in a cooperative manner (von Kries et al., 1991; Buhrmester et al., 1995)	ATTCAGGTGTAAAA TGCAGGTGTCCTT TGGGGGTGTAAAA	2 3 4	no yes yes
ATF	ATF sequences bind transcription factors of the ATF family, two of which were shown to be integral components of the nuclear matrix (Stein et al., 1991)	TGACGTCCATG	5	no
BEAF-32	Beaf-32 (boundary element-associated factor of 32 kDa molecular mass) binds to the locus boundary element scs' from <i>Drosophila</i> , which can insulate reporter genes from position effect variegation (Zhao et al., 1995)	CGATA	6	yes
topo-isomerase II	Topoisomerase II is a major component of the nuclear matrix and topoisomerase II binding sites have been found in most MARs to date (Sander and Hsieh, 1985; Boulikas, 1995)	CNNGYNGKTNYNY ASMATGCCGYWYATCRT GTNWAKATTNATNNR	7 8 9	yes no yes
unwinding sequence	Unwinding sequences such as AATATATTT cause DNA unwinding <i>in vivo</i> , which is important for MAR function (Bode et al., 1992)	AATATATTT	10	yes
SATB1	SATB1 is a protein isolated from human thymus tissue that selectively binds MAR sequences consisting of A's, T's and C's in one strand (Dickinson et al., 1992; Nagagomi et al., 1994)	TTCTAATATAT ATAATCTTC TTATTATTTA TATAAAAA AAGATTATATA TTTAATGAGATAATAA TATAATCTTC	11 12 13 14 15 16 17	no no yes yes no no no
bending DNA	Curved DNA regions are often found in or near MARs (Bode et al., 1995)	AAANNNNNNNAAA	18	yes
stem-loop	Stem loops create small single-stranded regions which are important for MAR function (Boulikas and Kong, 1993)			yes
oligo A/T tracts	These can create bent DNA that may attract protein complexes involved in topoisomerization,	AAAAAAA	19	yes

element	characteristics	sequence	SEQ ID NO:	reported in plant MARs
	recombination, transcription or replication (Travers, 1990)			

K: G or T, M: A or C, N: A, C, G or T, R: A or G, S: C or G, W: A or T, Y: C or T.

The present invention utilizes a subset of the features described in Table 1 in a novel artificial MAR.

5 The sequence of the 327 bp artificial MAR is given in SEQ ID NO:1. The artificial MAR was designed as a sequence flanked by *Bgl*II and *Bam*HI restriction sites, which are included in SEQ ID NO:1, but which are not critical to the function of the MAR. The functional portion of the 10 MAR comprises bp 11 to 309 of SEQ ID NO:1.

The following features are found in SEQ ID NO:1:

Feature	location (bp)
BglII	5-10
BEAF-32	11-15
SATBI	24-34
unwinding	28-36
topoisomerase II	44-59
ATF site	60-69
A/T tract	70-85
BEAF-32	86-90
stem-loop	93-101/117-124
unwinding	105-113
topoisomerase II	125-139
SATBI	149-159
unwinding	153-161
BEAF-32	164-168
SATBI	185-195
stem-loop	208-216/231-239
unwinding	219-227
A/T tract	241-253
topoisomerase II	268-283
curved (bending) DNA	284-294
ARBP site	295-309
BamHI	318-323

The 3' UTR, or 3' untranslated region, that is employed in constructs of the invention is one that confers efficient processing of the mRNA, maintains stability of the message and directs the addition of adenosine ribonucleotides to the 3' end of the transcribed mRNA sequence. The 3' UTR may be native with the promoter region, native with the structural gene, or may be derived from another source. Suitable 3' UTRs include, but are not limited to: the *per5* 3' UTR, and the 3' UTR of the nopaline synthase (*nos*) gene.

Example 1

Synthesis of artificial MAR

To construct the artificial MAR, six individual oligonucleotides were synthesized and assembled by PCR. The sequences for the six oligonucleotides, referred to hereinafter as MAR-A, MAR-B, MAR-C, MAR-D, MAR-E, and MAR-F, are given in the Sequence Listing as SEQ ID NOS: 20 through 25, respectively. A 15 bp overlap between adjacent oligonucleotides allowed assembly of the MAR by PCR, using the strategy shown in Figure 1.

The GeneAmp™ PCR Reagent Kit with AmpliTaq DNA Polymerase (Perkin Elmer, Norwalk, CT) was used for the DNA amplification. Twenty cycles of PCR (denaturation: 30 sec at 94°C; annealing: 60 sec at 52°C and extension: 5 60 sec at 70°C) with primers MAR-C and MAR-D were followed by 20 cycles of PCR with primers MAR-B and MAR-E, using the product from the first reaction as template for the second reaction. The 231 bp product of this reaction was purified from a low melting point agarose gel and used as 10 a template for 20 cycles of PCR with primers MAR-A and MAR-F. The 327 bp product from this reaction was subcloned into pCR2.1 using the TA Cloning™ Kit (Invitrogen, San Diego, CA) and the sequence was verified by sequencing.

15 A dimer consisting of two tandem copies of the *Bgl*III/*Bam*HI fragment was constructed in the *Bam*HI site of pBluescript™ SK- (Stratagene, La Jolla, CA). The sequence of the dimer is bp 5-630 of SEQ ID NO:26.

EXAMPLE 2

20 Binding of artificial MAR to nuclear scaffolds

A. Controls

Two DNA fragments of similar size and nucleotide composition as the artificial MAR were amplified from plant DNA to serve as controls in the binding assay.

25 These fragments were a 657 bp fragment from the 3' end of a maize gene Gpal (glyceraldehyde-3-phosphate dehydrogenase subunit A, GenBank accession number X15408, bases 4516 to 5173, Quigley et al., 1989), and a 488 bp fragment from the 5' flanking region of a 19 kD alpha 30 zein gene (GenBank accession number X05911, bases 339 to 827, Kriz et al., 1987). Table 2 compares the features present in the artificial MAR and control fragments.

Table 2

element	Artificial MAR dimer	Gpal control	zein control
ARBP sites	2	0	0

element	Artificial MAR dimer	Gpal control	zein control
ATF sites	2	0	0
BEAF-32 sites	6	1	1
topoisomerase II sites	6	0	0
unwinding sequence sites	8	0	0
SATB1 sites	6	0	0
bending DNA sites	yes	yes	yes
stem-loop sites	4	1	0
oligo A/T tracts	2	0	1
fragment size in binding assay	632	657	488
% A + T	71	63	66
strength of binding to nuclear scaffolds	+++++	-	-

B. Preparation of nuclei from maize leaves

Nuclei for use in isolating nuclear scaffolds were prepared from young maize leaves by adaptation of a published protocol (Hall *et al.*, 1991). Nuclei were counted and checked 5 for integrity by microscopic examination of DAPI stained aliquots. Only high quality nuclei were used to prepare nuclear scaffolds by lithium diiodosalicylate extraction.

For nuclei purification, young maize leaves from V4 stage plants (fourth or fifth leaf) were harvested with a razor 10 blade, washed and dried. After removing the midrib, leaves were frozen in liquid nitrogen, and ground to a fine powder with a mortar and pestle. The powdered leaf samples were transferred to a glass beaker, and 5 ml NIB1+PI (0.5 M hexylene glycol, 20 mM piperazine-N,N'-bis[2-ethanesulfonic acid] (PIPES), pH 6.5, 20 mM KCl, 7 mM 2-mercaptoethanol, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.4 % Triton X-100™ [Rohm & Haas Company, Philadelphia, PA], 0.05 mM spermine, 15 0.125 mM spermidine, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 µg/ml leupeptin, 1 µg/ml aprotinin) was added per gram of leaf tissue. The leaf extract was filtered sequentially through 1900, 520, 125, 85 and 40 mm filters at 20 4°C and filters were rinsed with 1 ml NIB1+PI per gram leaf to collect any nuclei that were trapped in the debris. Fifteen ml crude nuclear extract was loaded onto Percoll™ (Pharmacia 25 Biotech, Piscataway, NJ) gradients consisting of 7 ml 40% Percoll in NIB1 (0.5 M hexylene glycol, 20 mM PIPES, pH 6.5,

20 mM KCl, 7 mM 2-mercaptoethanol, 0.5 mM EDTA, 0.4 % Triton™ X-100, 0.05 mM spermine, 0.125 mM spermidine) and 5 ml 70% Percoll in NIB1. After centrifugation for 15 min at 500xg at 4°C the 40%/70% interface was collected with a sterile pasteur 5 pipette and added to 2 volumes NIB2 (0.5 M hexylene glycol, 20 mM PIPES, pH 6.5, 20 mM KCl, 7 mM 2-mercaptoethanol, 0.5 mM EDTA, 0.05 mM spermine, 0.125 mM spermidine), taking care to avoid the pellet and other debris. Nuclei were concentrated by centrifugation at 600xg for 10 min at 4°C.

10 The nuclear pellet was resuspended in 20 ml NIB2 and centrifuged as before. This step was repeated one more time to wash away traces of Percoll. Nuclei were counted using a hemacytometer and resuspended in NIB2+PI/50% glycerol (0.5 M hexylene glycol, 20 mM PIPES, pH 6.5, 20 mM KCl, 7 mM 2- mercaptoethanol, 0.5 mM EDTA, 0.05 mM spermine, 0.125 mM 15 spermidine, 50% glycerol, 1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin) at 20 million nuclei/ml. Nuclei were stored at -80°C until used for scaffold preparation.

20 C. Preparation of nuclear scaffolds

Frozen nuclei were thawed and washed with 10 ml of NIB3+PI (0.5 M hexylene glycol, 20 mM PIPES, pH 6.5, 20 mM KCl, 7 mM 2-mercaptoethanol, 0.05 mM spermine, 0.125 mM spermidine, 1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin) per 20 million nuclei. Nuclei were collected by centrifugation 25 at 600xg for 10 minutes, resuspended in 200 ml NIB3+PI in the presence of 1 mM CuSO₄, and incubated for 10 min at 42°C to stabilize the nuclei.

Histones were extracted by incubation in 10 ml HIB+PI 30 (20 mM HEPES, pH 7.4, 100 mM lithium acetate, 10 mM LIS (lithium diiodosalicylate), 0.1 % digitonin, 2 mM EDTA, 1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin) for 15 minutes at room temperature. The resulting nuclear halos were transferred to a centrifuge tube and pelleted at 4000xg for 10 minutes. 35 Halos were washed twice with 10 ml HWB (20 mM Tris, pH 8, 70

mM NaCl, 20 mM KCl, 7 mM 2-mercaptoethanol, 0.1 % digitonin, 0.05 mM spermine, 0.125 mM spermidine) and once with D/BB+PI (HWB + 10 mM MgCl₂, 1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin) to remove LIS. If halos did not pellet well, 5 subsequent centrifugation steps were done at 6000xg using a slow brake setting. The quality of the halos was verified by SDS-PAGE gel to ensure that more than 95% of the histones were removed in the extraction procedure

Washed nuclear halos were resuspended in 400 µl D/BB+PI 10 and 200 units of restriction enzymes (100 u each of *Eco*RI and *Hind*III) were added, and incubated at 37°C for 2-3 hours on a rocking platform to keep the halos from settling. The restriction enzymes removed more than 70% of the nuclear DNA, producing nuclear scaffolds. Scaffolds were pelleted at 300xg 15 and washed with HWB+PI (HWB + 1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin). Nuclear scaffolds were resuspended in 400 µl HWB+PI and separated into 100 µl aliquots (containing 5 million nuclear equivalents).

20 D. Binding of artificial MAR to nuclear scaffolds

100 µl aliquots of scaffolds in HWB+PI were incubated with probe and *E. coli* competitor DNA at 37°C for 2-3 hours in siliconized microfuge tubes on a rocking platform shaker. After incubation, the supernatant fraction (containing unbound 25 DNA fragments) and pellet fraction (containing scaffolds and bound DNA fragments) were separated via centrifugation in a horizontal microfuge at 3000xg for 5 min. The pellet was washed once with 200 µl HWB to remove proteinase inhibitors, resuspended in 100 µl lysis buffer (10 mM Tris, pH 8, 10 mM 30 EDTA, 0.5 % SDS, 0.5 mg/ml Proteinase K) and incubated overnight at room temperature.

Equal fractions of the pellet and supernatant were separated on a 0.9% agarose gel, which was subsequently fixed, by soaking in 7% TCA for 20 min, dried and exposed to X-ray

film at room temperature and/or storage phosphor screens for the PhosphoImager™ SI (Molecular Dynamics, Sunnyvale, CA).

Plasmids containing the artificial MAR monomer or dimer or the control Gpal or zein sequences were digested with restriction enzymes that generate 5' overhang ends. The Klenow subunit of DNA polymerase I was used to fill the overhang with [α -³²P]dCTP (Amersham Life Science, Arlington Heights, IL). The end-labeled DNA fragments were used as probes in the binding assay, i.e. the fragments were incubated with purified maize nuclear scaffolds in the presence of unlabeled *E.coli* competitor DNA and the relative binding of the inserts was determined. Relative amounts of nuclei, probe and unlabeled *E. coli* competitor DNA used in the binding assay were optimized to obtain maximal discrimination between strongly and weakly binding MARs. The optimal relative amounts were 2, 5 or 10 µg of unlabeled *E. coli* competitor DNA, 5 million nuclear equivalents of nuclear scaffolds, and 1 fmole of digested and labeled plasmid per assay.

The artificial MAR dimer bound very strongly to the nuclear scaffold preparation, even in the presence of high levels of competitor DNA. The monomer MAR also bound to nuclear scaffold preparations, albeit at a lower affinity. Neither control sequence was retained in the pellet fraction, even though they were similar to the artificial MARs in size and relative AT content. This suggests that the elements included in the artificial MAR facilitate binding.

EXAMPLE 3

EVALUATION OF THE ARTIFICIAL MAR IN RICE

A. Rice Transformation Vectors

pGOS2-hpt (SEQ ID NO:27) is a rice transformation vector containing a hygromycin selectable marker driven by the 35S promoter and a GOS2/GUS/nos cassette (GOS2 transcription initiation region/GUS structural gene/nos 3' untranslated region). The GOS2 transcription initiation region in this construct is comprised of 1010 bp of promoter and 170 bp of

untranslated 5' leader interrupted by a 1100 bp intron (de Pater et al., 1992).

pArGOS2Af-hpt (SEQ ID NO:28) is a rice transformation vector identical to pGOS2-hpt except that it has the MAR dimer of SEQ ID NO:29 positioned 5' to the GOS2 transcription initiation region and the MAR dimer of SEQ ID NO:26 positioned 3' to the nos 3' UTR.

A schematic representation of the ArGOS2Af construct is shown in FIG 2.

10 B. Transformation of Rice

For initiation of embryogenic callus, mature seeds of a *Japonica* cultivar, Taipei 309, were dehusked and surface-sterilized in 70% ethanol for 5-7 min. followed by soaking 30-45 min in 25% commercial bleach (2.6% sodium hypochlorite) with 0.02% Tween™ 20 (ICI Americas, Inc.) under vacuum. The seeds were then rinsed 5 times in sterile distilled water and placed on filter paper before transferring to induction media (NB). The NB medium consisted of N6 macro elements (Chu, 1978), B5 micro elements and vitamins (Gamborg et al., 1968), 300 mg/l casein hydrolysate, 500 mg/l L-proline, 500 mg/l L-glutamine, 30 g/l sucrose, 2 mg/l 2,4-dichloro-phenoxyacetic acid (2,4-D), and 2.5 g/l Gelrite™ (Merck & Co., Rawhay, NJ) with the pH adjusted to 5.8. The mature seed cultured on induction media were incubated in the dark at 28° C for three weeks. Primary callus induced from the scutellar region of mature embryo was transferred to fresh NB medium for further maintenance and thereafter maintained on a two week subculture period.

To prepare DNA for blasting, about 140 µg of plasmid DNA (pGOS2-hpt or pArGOSAf-hpt) was precipitated onto 60 mg of gold particles. The plasmid DNA was precipitated onto 1.5-3.0 micron (Aldrich Chemical Co., Milwaukee, WI) or 1.0 micron gold particles (Bio-Rad Laboratories, Hercules, CA). The precipitation mixture included 60 mg of pre-washed gold particles, 300 µl of water/DNA (140 µg), 74 µl of 2.5 M CaCl₂, and 30 µl of 0.1 M spermidine. After adding the components in

the above order, the mixture was vortexed immediately, and allowed to settle for 2-3 min. The supernatant was pipetted off and discarded. The DNA-coated gold particles were resuspended in 1 ml of 100% ethanol and diluted to 17.5 mg 5 DNA/7.5 mg gold per ml of ethanol for use in blasting experiments.

For helium blasting, actively growing embryogenic callus cultures, 2-4 mm in size, were subjected to a high osmoticum treatment by placing callus on NB medium with 0.2 M mannitol 10 and 0.2 M sorbitol (Vain et al., 1993) for 4 hr before helium blasting. Following osmoticum treatment, callus cultures were transferred to blasting medium (NB+2% agar) and covered with a stainless steel screen (230 micron). Helium blasting involved accelerating the suspended DNA-coated gold particles towards 15 and into the prepared tissue targets. The device used was an earlier prototype to the one described in US Patent No. 5,141,131, which is incorporated herein by reference, although both function in a similar manner. The callus cultures were blasted at different helium pressures (1750-2,250 psi) one to 20 three times per target. After blasting, callus was transferred back to the high osmotic media overnight before placing on selection medium, which consisted of NB medium with 30 mg/l hygromycin. After 2 weeks, the cultures were transferred to fresh selection medium with higher 25 concentrations of selection agent, i.e., NB+50 mg/l hygromycin (Li et al., 1993).

Compact, white-yellow, embryogenic callus cultures, recovered on NB+50 mg/l hygromycin, were regenerated by transferring to pre-regeneration (PR) medium+50 mg/l 30 hygromycin. PR medium consisted of NB medium with 2 mg/l benzyl aminopurine (BAP), 1 mg/l naphthalene acetic acid (NAA), and 5 mg/l abscisic acid (ABA). After 2 weeks of culture in the dark, they were transferred to regeneration (RN) medium. The composition of RN medium is NB medium with 35 3 mg/l BAP, and 0.5 mg/l NAA. The cultures on RN medium were incubated for 2 weeks at 28° C under high fluorescent light

(325-ft-candles). The plantlets with 2 cm shoots were transferred to 1/2 MS medium (Murashige and Skoog, 1962) with 1/2 B5 vitamins, 10 g/l sucrose, 0.05 mg/l NAA, 50 mg/l hygromycin and 2.5 g/l Gelrite™ adjusted to pH 5.8 in GA7 vessels (Magenta Corp., Chicago, IL). When plantlets were established with well-developed root systems, they were transferred to soil [1 part Metro-Mix 360 (Scotts-Sierra Horticultural Products Co., Marysville, OH) and 1 part top soil] and raised in a growth chamber (29/24°C day/night cycle, 5 50-60% humidity, 12 h photoperiod) until they reached a height of 60 cm, at which point 2 leaves were harvested for quantitative GUS analysis, and the plants were transferred to the greenhouse to grow to maturity.

15 C. Southern analyses

Southern analysis was used to identify primary regenerate (R_0) rice lines lines that contained intact copies of the specific gene construct.

A DNA probe specific for the coding region of the β -glucuronidase (GUS) gene construct was gel purified with the Qiaex II DNA purification kit (Qiagen Inc., Chatsworth, CA). Radiolabeled probe was prepared using the Ready-To-Go™ DNA labeling beads (Pharmacia LKB, Piscataway, NJ) with 50 microcuries of [$\alpha^{32}\text{P}$]dCTP (Amersham Life Science, Arlington Heights, IL).

Leaf material from R_0 rice plants was harvested from two representatives from each line. Genomic DNA from the R_0 plants was prepared from lyophilized tissue as described by Saghai-Marof et al. (1984).

30 Four micrograms of rice DNA was digested with restriction enzyme to release the intact gene construct using conditions suggested by the manufacturer (Bethesda Research Laboratory, Gaithersburg, MD) and separated by agarose gel electrophoresis. The DNA was blotted onto nylon membranes as 35 described by Southern (1975, 1989). Radiolabeled probe DNA was hybridized to the genomic DNA on the blots using 50 ml of

minimal hybridization buffer (10% polyethylene glycol, 7% sodium dodecyl sulfate, 0.6x SSC, 10 mM sodium phosphate, 5 mM EDTA and 100 mg/ml denatured salmon sperm DNA) heated to 60°C and mixed with the denatured radiolabeled probe prior to 5 being added to the blots for overnight hybridization at 60°C. The blots were washed at 60°C in 0.25X SSC and 0.2% SDS for 45 minutes, blotted dry and exposed to XAR-5 film with two intensifying screens overnight.

Southern analysis was conducted on seventy ArGOS2Af R₀ 10 rice lines. The DNA from the R₀ plants was digested with the restriction enzyme *Xba*I which, if the intact gene construct is present, should result in a 5.7 kb hybridization product when radiolabeled with a probe specific for the GUS coding region. The 5.7 kb fragment should consist of the artificial MAR in 15 the reverse orientation, the GOS2 promoter, the GUS coding region, the nos 3' UTR and the artificial MAR in the forward orientation. The expected 5.7 kb hybridization product was detected in twenty-five of the seventy rice lines. All of the twenty-five lines had multiple hybridization products and two 20 of the lines had identical complex hybridization patterns indicating that they are probably from the same transformation event.

The non-Mar control lines, GOS2, were also analyzed by Southern analysis. The DNA from forty-eight GOS2 R₀ lines was 25 digested with the restriction enzymes *Eco*RI and *Xba*I which, if the intact gene is present, should result in a 4.4 kb hybridization product when radiolabeled with a probe specific for the GUS coding region. The 4.4 kb fragment would include 1.6 kb of the GOS2 promoter, the GUS coding region, the nos 30 3' UTR and the 35T promoter (the promoter used to drive the selectable marker gene). The expected 4.4 kb hybridization product was detected in twenty-eight of the forty-eight GOS2 lines. Two of the lines had identical hybridization patterns and must have resulted from the same transformation event. 35 Two of the lines contained genetic chimeras.

D. GUS analysis

Analysis of rice was performed on young leaves of primary transformants, after plants had been grown 6-8 weeks in an environmentally controlled growth chamber and had reached a height of about 60 cm. Two independently regenerated rice plants were analyzed per transformation event. Individual transformants were analyzed by Southern blots to verify the presence of an intact copy of the transgene and determine whether each event displayed unique hybridization patterns, indicating independent transformation events. Plants lacking a complete copy of the transgene, chimeric events, or duplicated integration events were not included in the analysis.

Results of the analysis are reported in FIGS 4 and 5. In FIG 4, error bars represent the standard deviation between the plants for each transformation event. Two samples were independently processed for each plant. In general, the level of expression of GUS in independent rice plants from each transformation event was similar (as demonstrated by the standard deviation of the results shown in FIG 4).

FIG 5 reports the percent of transformation events expressing GUS in the indicated ranges.

EVALUATION OF ARTIFICIAL MAR IN ARABIDOPSIS

A. *Arabidopsis* Transformation Vectors

Act2/GUS/nos (Act2 transcription initiation region/GUS structural gene/nos 3' UTR) constructs were made for testing in a dicot system (*Arabidopsis*). Three vectors were made:

pAct2-bin (SEQ ID NO:30) is a binary vector containing a Act2/GUS/nos cassette, 19S/NPTII/orf25polyA as a selectable marker, and 35S/GFP/nos as an independent reporter gene.

pArAct2Af-bin (SEQ ID NO:31) is identical to pAct2-bin except that it has the MAR dimer of SEQ ID NO:29 positioned 5' to the Act2 transcription initiation region, and the MAR dimer of SEQ ID NO:26 positioned 3' to the nos 3' UTR.

pAfAct2Af-bin (SEQ ID NO:32) is identical to pAct2-bin except that it has the MAR dimer of SEQ ID NO:26 positioned 5'

to the Act2 transcription initiation region the MAR dimer of SEQ ID NO: 33 positioned 3' to the nos 3' UTR.

These vectors enabled testing of two orientations of the artificial MAR dimer in *Arabidopsis*. A schematic of the 5 pArAct2Af-bin and pAfAct2Af-bin constructs is shown in Figure 3.

B. *Arabidopsis* transformation

Arabidopsis transformation was performed according to a 10 protocol provided by Pam Green (van Hoof and Green 1996), which is an adaptation from protocols by Nicole Bechtold (Bechtold et al., 1993), Andrew Bent (Bent et al., 1994) and Takashi Araki (personal communication).

Seeds of ecotype Columbia were planted in 4 inch square 15 pots, covered with window screen mesh, and grown under conditions of 16 hours light/ 8 hours dark at 22°C, fertilizing by subirrigation once a week. The fertilizer consisted of 5 mM KNO₃, 2.5 mM KPO₄ (pH 5.5), 2 mM MgSO₄, 2 mM Ca(NO₃)₂, 0.05 mM Fe•EDTA, 0.07 mM boric acid, 0.014 mM MnCl₂, 0.005 mM CuSO₄, 20 0.001 mM ZnSO₄, 0.0002 mM NaMoO₄, and 0.01 mM NaCl. Plants were thinned to 4 plants per pot and grown until several bolts emerged. When plants were ready to transform, the above soil parts were submerged in infiltration medium (2.2 g/l MS salts, 1X B5 vitamins, 50 g/l sucrose, 2.5 mM MES, pH 5.7, 0.044 M 25 benzylaminopurine, 200 ml/l Silwet L-77™ [Osi Specialties, Inc.] containing *Agrobacterium* cells, placed inside a vacuum desiccator under a vacuum of 400 mm Hg (about 17 inches) for 5 minutes. After quickly releasing the vacuum, pots were drained and placed on their sides in a tray covered with 30 plastic wrap to maintain humidity for 24 hours. The next day the pots were uncovered and set upright. Plants were staked individually and after 2 weeks watering was gradually reduced to allow plants to dry out. Seeds were harvested from each plant individually.

35 For selection of transformation events, 1-10 mg seeds per plant were surface sterilized by soaking in 10% bleach for 7

minutes while mixing vigorously, followed by three rinses in sterile water, and placed in a flask containing *Arabidopsis* germination medium (MS salts, MS vitamins, 10% sucrose, 2.5 mM 2-[N-morpholino]ethanesulfonic acid [MES], 30 mg/l kanamycin, 5 50 mg/l vancomycin and 0.1 % Bacto™-Agar [Difco Laboratories, Detroit, MI]). After shaking in continuous light at 90 rpm for 3 days, seeds germinated, and transformants were isolated as green seedlings between 7 and 12 days after germination. Nontransformed seeds produced small bleached seedlings.

10 Transformants were transferred to solid medium (MS salts, B5 vitamins, 10% sucrose, 2.5 mM MES, 15 g/l Phytagar™ [Gibco BRL, Gaithersburg, MD], 30 ml/l kanamycin, 50 mg/l vancomycin) in plates for further selection. After one to two weeks, true transformants were transferred to GA7 vessels (MS salts, B5 15 vitamins, 0.3 % sucrose, 2.5 mM MES) for one to two weeks prior to planting in soil for production of T1 seed.

C. Southern analyses

Southern analysis was used to identify primary regenerate 20 T2 *Arabidopsis* lines that contained intact copies of the specific gene construct.

Pooled samples of *Arabidopsis* leaf tissue were powdered in liquid nitrogen. The ground tissue was then incubated for three minutes in 500 µl 2X extraction buffer (2% CTAB, 100 mM 25 Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl and 2% 2-mercaptoethanol) at 65°C. Five hundred µl of chloroform/octanol (24:1) was added, the samples were shaken for two minutes, and then spun at 14,000 ' g in a microcentrifuge and the supernatant was removed. The 30 chloroform/octanol extraction was repeated. One ml of precipitation buffer (1% CTAB, 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, and 1% 2-mercaptoethanol) was added to the supernatant and then incubated at room temperature for 60 minutes. The DNA was pelleted by centrifugation at 3500g for 5 minutes in a 35 microcentrifuge. The pellet was drained and resuspended in 200 µl 1.0 M NH₄OAc. One hundred µl 7.5 M NH₄OAc and 1 ml

isopropanol were added, the samples incubated on ice for 5 minutes and then centrifuged at 14,000 ' g for 5 minutes. The pellet was drained and resuspended in 200 µl TE. 100 µl 7.5 M NH₄OAc and 1 ml isopropanol were added and incubated on ice for 5 minutes then centrifuged at 14,000 ' g for 5 minutes. The pellet was drained and rinsed with 70% ethanol and dried in a Speed Vac (Savant Instruments Inc., Farmingdale, NY). The dried pellet was resuspended in 20 µl TE (10 mM TRIS, 1 mM EDTA, pH 8.0).

10 Southern analysis was conducted on 29 ArAct2Af T2 lines and 24 AfAct2Af T2 lines. One microgram of DNA from the ArAct2Af plants was digested with the restriction enzyme *Xba*I using conditions suggested by the manufacturer (Bethesda Research Laboratory, Gaithersburg, MD) and separated by 15 agarose gel electrophoresis, which should result in a 4.6 kb hybridization product, if the gene construct is intact, when radiolabeled with a probe specific for the GUS coding region. Similar to the ArGOS2Af rice plants, the 4.6 kb fragment should consist of the artificial MAR in the reverse orientation, the Act2 promoter, the GUS coding region, the nos 20 3' UTR and the artificial MAR in the forward orientation. The DNA was blotted onto nylon membranes as described by Southern (1975, 1989). Radiolabeled probe DNA was hybridized to the genomic DNA on the blots using 50 ml of minimal hybridization 25 buffer (10% polyethylene glycol, 7% sodium dodecyl sulfate, 0.6x SSC, 10 mM sodium phosphate, 5 mM EDTA and 100 mg/ml denatured salmon sperm DNA) was heated to 60°C and mixed with the denatured radiolabeled probe prior to being added to the blots for overnight hybridization at 60°C. The blots were 30 washed at 60°C in 0.25X SSC and 0.2% SDS for 45 minutes, blotted dry and exposed to XAR-5 film with two intensifying screens overnight.

The expected 4.6 kb hybridization product was detected in twenty-six of the twenty-nine ArAct2Af lines. A second 35 Southern blot was generated to determine the copy number of the ArAct2Af construct. The ArAct2Af DNA was digested with

the restriction enzymes *Sst*I and *Xho*I which cut the construct near the right and left borders of the T DNA. The blots were radiolabeled with probes specific for the artificial MAR and the DNA from the left border to the *Xho*I site 800 bp downstream. A single copy of the ArAct2Af construct will have three hybridization products: two fragments of unknown size, consisting of the left and right border DNA, and the 8.9 kb fragment which is the DNA internal to the borders. Twenty-two of the twenty-nine lines had three or fewer hybridization products, indicating that a single copy of the ArAct2Af construct was present.

The DNA from the AfAct2Af plants was digested with the restriction enzyme *Xba*I which, if the construct remains intact, should result in a 5.7 kb hybridization product when radiolabeled with a probe specific for the GUS coding region. The 5.7 kb fragment should consist of the artificial MAR in the forward orientation, the Act2 promoter, the GUS coding region, the nos 3' UTR and the artificial MAR in the forward orientation. It also includes the green fluorescent protein coding region and the nos 3' UTR. The expected 5.7 kb hybridization product was detected in all of the twenty-four AfAct2Af lines.

A second Southern blot was generated to determine the copy number of the AfAct2Af construct. The AfAct2Af DNA was digested with the restriction enzymes *Sst*I and *Xho*I which cut the construct near the right and left borders of the T DNA. The blots were radiolabeled with probes specific for the artificial MAR and the DNA from the left border to the *Xho*I site 800 bp downstream. A single copy of the AfAct2Af construct will have three hybridization products: two fragments of unknown size, consisting of the left and right border DNA and the 8.9 kb fragment which is the DNA internal to the borders. Fifteen of the twenty-four lines had three or fewer hybridization products indicating that a single copy of the AfAct2Af construct was present.

The non-Mar control lines, Actbin, were also analyzed by Southern analysis. The DNA from thirty-four Actbin T2 lines was digested with the restriction enzyme *Pst*I which should, if the construct remains intact, result in a 3.4 kb hybridization product when radiolabeled with a probe specific for the GUS coding region. The 3.4 kb fragment should consist of the *Act2* promoter, the GUS coding region, and the nos 3' UTR. The expected 3.4 kb hybridization product was detected in thirty of the thirty-four Actbin lines. A second Southern blot was generated to determine the copy number of the Actbin construct. The DNA from the Actbin lines was digested with the restriction enzymes *Sst*I and *Xho*I which cut the construct near the right and left borders of the T DNA. The blots were radiolabeled with probes specific for the nos 3'UTR and the DNA from the left border to the *Xho*I site 800 bp downstream. A single copy of the Actbin construct will have three hybridization products: two fragments of unknown size, consisting of the left and right border DNA and the 8.2 kb fragment which is the DNA internal to the borders. Nine of the thirty-four lines had three or fewer hybridization products indicating that a single copy of the Actbin construct was present.

D. GUS analysis

For growing *Arabidopsis*, T2 seed was germinated *in vitro* on MS medium containing 90 mg/l kanamycin and 3 week old kanamycin resistant seedlings were harvested. Two batches of 30 seedlings per transformation event were used for GUS analysis and additional seedlings were used to extract DNA.

For analyses of GUS activity, leaf samples were powdered in liquid nitrogen and samples of approximately 400 ml of tissue were placed in microfuge tubes. Two independent samples from each leaf sample was processed. The tissue was either stored at -70°C or extracted immediately. GUS was extracted by mixing the powdered tissue with GUS lysis buffer (Jefferson *et al.*, 1987) modified by the addition of 1%

polyvinylpolypyrrolidone (hydrated in the buffer for at least one hour) and 20% glycerol. After incubation on ice for at least 10 min, the samples were centrifuged at 16,000 ' g for 10 min. The supernatants were recovered and centrifuged a second time as described above. The supernatants were recovered and frozen on dry ice and stored at -70°C. Experiments showed that GUS activity was stable for at least 4 freeze-thaw cycles when stored in the buffer described above (W.M. Ainley, unpublished). GUS activity was measured using a GUS-Light™ kit (Tropix, Inc., Bedford, MA). Five ml samples of undiluted extract or of extract diluted so that the luminescence was within the range measured by the luminometer was added to 195 ml of the GUS-Light™ Reaction Buffer. Luminescence was integrated for 5 sec after a 5 sec delay. Protein was measured with the assay developed by Bradford (1976) using human serum albumin as the standard. GUS activity was normalized between experiments using a GUS standard obtained from SIGMA. The amount of plant protein in the standards and plant samples was the same; protein was adjusted where necessary using extracts of nontransformed plants.

FIGS 6 and 7 report the results the GUS analysis. FIG 6 reports expression observed for independent transformation events expressing either the base, non-MAR construct (Actbin) or the base construct flanked by the artificial MAR in the indicated orientation (ArAct2Af or Af-Act-Af). Standard deviations indicate the variance between plants harvested from different plates. Arrows indicate the relative orientations of the MARs.

E.

In FIG 7, percent of transformation events expressing GUS in the indicated ranges is shown.

E. Characterization of transgenic plants expressing the reporter gene constructs

T2 *Arabidopsis* plants were analyzed. The plants were evaluated for segregation of kanamycin resistance to determine

insert copy number. A few events did not fall into either one or two insert categories as determined by chi square analyses. Of the remaining plants, 72% had a single insert. Most of the inserts had multiple copies of the transgene. Although 5 the *Arabidopsis* transformation events were grown under controlled environments, there were some occasional differences in the relative growth of plants between the duplicate plates. In those cases, the events were either not used or were grown again. The coefficient of variance of the 10 expression determined from duplicate plates range of between 1 and 46%, with 89 percent below 20 percent (Figure 6).

Generally, all *Arabidopsis* transformation events tested for a construct were grown and analyzed together. Although the growth of the plants was under controlled environments, 15 the possibility remained that the differences observed between transformation events expressing the constructs was due to environmental differences occurring between different experiments. To eliminate this possibility, three of the highest expressing events from the three sets of *Arabidopsis* 20 transformation events were grown together and reanalyzed. This analyses confirmed the earlier differences between the sets of transformation events.

SUMMARY OF EFFECTS OF THE ARTIFICIAL MAR ON TRANSGENE
25 EXPRESSION

In both rice and *Arabidopsis*, the average expression level of transformation events expressing the MAR-containing constructs expressed at a higher level than those lacking the 30 MAR elements (Table 3). In *Arabidopsis*, orientation of the MARs tested influenced the level of expression. Plants containing constructs in which the MARs were in the same orientation on either side of the GUS gene construct expressed GUS at higher levels than plants containing constructs in 35 which the MARs were oriented in opposite orientations.

There appeared to be a proportionally higher expression level in the higher expressors than in the lower expressors in

each set of transformation events. To document this, expression of the upper quartile of the expressors was compared (Table 3). Based on this data, the upper quartile of the transformation events expressing the AfAct2Af construct produce GUS protein at levels 5.6 times higher than in the upper quartile of events expressing the Actbin construct. The artificial MARs in opposite orientations in both rice and *Arabidopsis* enhance expression approximately two-fold over the respective constructs lacking MARs.

10

Table 3

Comparison of the expression of transformation events expressing either the base constructs or base constructs flanked by the artificial MARs.

	CONSTRUCT				
	GOS2	ArGOS2Af	ActBin	Ar-Act-Af	Af-Act-Af
species transformed	rice	rice	<i>Arabidopsis</i>	<i>Arabidopsis</i>	<i>Arabidopsis</i>
method of transformation	particle bombardment	particle bombardment	<i>Agrobacterium</i>	<i>Agrobacterium</i>	<i>Agrobacterium</i>
number of transformation events analyzed	28	25	28	21	22
average	357	713	2,313	3,673	9,042
median	95	317	2,205	2,772	5,730
upper quartile range	258-3,076	1,674-2,661	3,047-5,311	6,240-10,614	15,595-27,179
upper quartile average	1,169	2,099	4,056	7,812	22,688
relative average expression	1.0	2.0	1.0	1.6	4.0
relative upper average quartile expression	1.0	1.8	1.0	1.9	5.6

averages for the upper quartile of transformation event expression levels were statistically different (p is less than or equal to 0.05) based on t-test analyses. (Gopal K. Kanji 1995)

Previous published studies (reviewed by Holmes-Davis and Comai, 1998) have shown that, with the exception of one MAR, all MARs tested to date enhance expression of reporter genes. This study represents the first report that a MAR constructed using elements found preferentially in MARs can enhance expression in plant species representing both monocotyledonous and dicotyledonous plants.

REFERENCES

- 5 Allen, G.C., Hall, G., Jr., Michalowski, S., Newman, W.,
 Spiker, S., Weissinger, A.K. and Thompson, W.F. (1996) High-
 level transgene expression in plant cells: Effects of a strong
 scaffold attachment region from tobacco. *Plant Cell* **8**, 899-
 913.
- 10 Allen, G.C., Hall, G.E., Jr., Childs, L.C., Weissinger, A.K.,
 Spiker, S. and Thompson, W.F. (1993) Scaffold attachment
 regions increase reporter gene expression in stably
 transformed plant cells. *Plant Cell* **5**, 603-613.
- 15 An, Y.-Q., McDowell, J.M., Huang, S., McKinney, E.C.,
 Chambliss, S. and Meagher, R.B. (1996) Strong constitutive
 expression of the *Arabidopsis ACT2/ACT8* actin subclass in
 vegetative tissues. *Plant J.* **10**, 107-121.
- 20 Avramova, Z. and Bennetzen, J.L. (1993) Isolation of matrices
 from maize leaf nuclei: identification of a matrix-binding
 site adjacent to the Adh 1 gene. *Plant Mol. Biol.* **22**, 1135-
 1143.
- 25 Avramova, Z., SanMiguel, P., Georgieva, E. and Bennetzen,
 J.L. (1995) Matrix attachment regions and transcribed
 sequences within a long chromosomal continuum containing maize
 Adh1. *Plant Cell* **7**, 1667-1680.
- 30 Bode, J., Kohwi, Y., Dickinson, L., Joh, Y., Klehr, D.,
 Mielke, C. and Kohwi-Shigematsu, T. (1992) Biological
 significance of unwinding capability of nuclear matrix-
 associated DNAs. *Science* **255**, 195-197.

35

Bode, J., Schlake, T., Rios-Ramirez, M., Mielke, C., Stengert, M., Kay, V. and Klehr-Wirth, D. (1995) Scaffold/matrix-attached regions: Structural properties creating transcriptionally active loci. *Int. Rev. of Cytol.* **162A**, 389-

5 454.

Bonifer, C., Vidal, M., Grosveld, F. and Sippel, A.E. (1990) Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice.

10 *EMBO J.* **9**, 2843-2848.

Boulikas, T. (1995) Chromatin domains and the prediction of MAR sequences. *Int. Rev. of Cytol.* **162A**, 279-388.

15 Boulikas, T. and Kong, C.F. (1993) Multitude of inverted repeats characterize a class of anchorage sites of chromatin loops to the nuclear matrix. *J. Cell. Biochem.* **53**, 1-12.

20 Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

25 Breyne, P., Van Montagu, M., Depicker, A. and Gheysen, G. (1992) Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco. *Plant Cell* **4**, 463-471.

30 Breyne, P., Van Montagu, M. and Gheysen, G. (1994) The role of scaffold attachment regions in the structural and functional organization of plant chromatin. *Transgenic Res.* **3**, 195-202.

35 Buhrmester, H., von Kries, J.P. and Stratling, W.H. (1995) Nuclear matrix protein ARBP recognizes a novel DNA sequence motif with high affinity. *Biochemistry* **34**, 4108-4117.

De Pater, B.S., van der Mark, F., Rueb, S., Katagiri, F.,
Chua, N-H., Schilperoort, R.A. and Hensgens, L.A.M. (1992)

The promoter of the rice gene GOS2 is active in various
5 different monocot tissues and binds rice nuclear factor ASF-1.
Plant J. **2**, 837-844.

Dennis, E.S., Gerlach, W.L., Pryor, A.J., Bennetzen, J.L.,
Inglis, A., Llewellyn, D., Sachs, M.M., Ferl, R.J. and

10 Peacock, W.J. (1984) Molecular analysis of the alcohol
dehydrogenase (*Adh1*) gene of maize. *Nucl. Acids Res.* **12**, 3983-
4000.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and
15 Goodman, H.M. (1982) Nopaline synthase: Transcript mapping
and DNA sequence. *J. Mol. Appl. Genet.* **1**, 561-573.

Dickinson, L.A., Joh, T., Kohwi, Y. and Kohwi-Shigematsu, T.
(1992) A tissue-specific MAR/SAR DNA-binding protein with
20 unusual binding site recognition. *Cell* **70**, 631-645.

Dillon, N. and Grosfeld, F. (1994) Chromatin domains as
potential units of eukaryotic gene function. *Curr Opinion Gen
Dev* **4**, 260-264.

25 Eggert, H. and Jack, R.S. (1991) An ectopic copy of the
Drosophila ftz associated SAR neither reorganizes local
chromatin structure nor hinders elution of a chromatin
fragment from isolated nuclei. *EMBO J.* **10**, 1237-1243.

30 Forrester, W.C., van Genderen, C., Jenuwein, T. and
Grosschedl, R. (1994) Dependence of enhancer-mediated
transcription of the immunoglobulin m gene on nuclear matrix
attachment regions. *Science* **265**, 1221-1225.

- 14 Franck, A., Guille, H., Jonard, G., Richards, K. and Hirth, L. (1980) Nucleotide sequence of cauliflower mosaic virus DNA. *Cell* **21**, 285-294.
- 5 Gritz, L. and Davies, J. (1983) Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. *Gene* **25**, 179-188.
- 10 Hall, G., Jr., Allen, G.C., Loer, D.S., Thompson, W.F. and Spiker, S. (1991) Nuclear scaffolds and scaffold-attachment regions in higher plants. *Proc. Natl Acad. Sci. USA* **88**, 9320-9324.
- 15 Jackson, D.A. (1995) Nuclear organization: uniting replication foci, chromatin domains and chromosome structure. *Bioessays* **17**, 587-591.
- 20 Jefferson, R.A. (1987) Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol. Biol. Rep.* **5**, 387-405.
- 25 Jefferson, R.A., Burgess, S.M. and Hirsh, D. (1986) b-Glucuronidase from *Escherichia coli* as a gene-fusion marker. *Proc. Natl Acad. Sci. USA* **83**, 8447-8451.
- 30 Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: b-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901-3907.
- 35 Kanji, G. K., (1995) *100 Statistical Tests*, Sage Publications Inc., Thousand Oaks CA
- 40 Kriz, A.L., Boston, R.S. and Larkins, B.A. (1987) Structural and transcriptional analysis of DNA sequences flanking genes

that encode 19 kilodalton zeins. *Mol. Gen. Genet.* **207**(1), 90-98.

5 **Lewin, B.** (1994) Chromatin and gene expression: Constant questions, but changing answers. *Cell* **79**, 397-406.

McKnight, R.A., Shamay, A., Sankaran, L., Wall, R.J. and Hennighausen, L. (1992) Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in 10 transgenic mice. *Proc. Natl Acad. Sci. USA* **89**, 6943-6947.

15 **Mlynarova, L., Jansen, R.C., Conner, A.J., Stiekema, W.J. and Nap, J-P.** (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. *Plant Cell* **7**, 599-609.

20 **Mlynarova, L., Loonen, A., Heldens, J., Jansen, R.C., Keizer, P., Stiekema, W.J. and Nap, J-P.** (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-attachment region. *Plant Cell* **6**, 417-426.

25 **Morrison, D.A., Trombe, M.C., Hayden, M.K., Waszak, G.A. and Chen, J.** (1984) Isolation of transformation-deficient Streptococcus pneumoniae mutants defective in control of competence, using insertion-duplication mutagenesis with the erythromycin resistance determinant of pAMb1. *J. Bacteriol.* **159**, 870-876.

30 **Mullineaux, P.M., Donson, J., Morris-Krsinicich, B.A.M., Boulton, M.I. and Davies, J.W.** (1984) The nucleotide sequence of maize streak virus DNA. *EMBO J.* **3**, 3063-3068.

Nagagomi, K., Kohwi, Y., Dickinson, L.A. and Kohwi-Shigematsu, T. (1994) A novel DNA binding motif in the nuclear

matrix attachment DNA-binding protein SATB1. *Mol. Cell. Biol.* **14**, 1852-1860.

Neznanov, N., Kohwi-Shigematsu, T. and Oshima, R.G. (1996)

5 Contrasting effects of the SATB1 core nuclear matrix attachment region and flanking sequences of the keratin 18 gene in transgenic mice. *Mol. Biol. Cell* **7**, 541-552.

Quigley, F., Brinkmann, H., Martin, W.F. and Cerff, R. (1989)

10 Strong functional GC pressure in a light-regulated maize gene encoding subunit GAPA of chloroplast glyceraldehyde-3-phosphate dehydrogenase: implications for the evolution of GAPA pseudogenes. *J. Molec. Evol.* **29(5)**, 412-421.

15 **Sander, M. and Hsieh, T-S.** (1985) Drosophila topoisomerase II double-strand cleavage: Analysis of DNA sequence homology at the cleavage site. *Nucl. Acids Res.* **13**, 1057-1072.

Schöffl, F., Schroder, G., Kliem, M. and Rieping, M. (1993)

20 An SAR sequence containing 395 bp DNA fragment mediates enhanced, gene-dosage-correlated expression of a chimaeric heat shock gene in transgenic tobacco plants. *Transgenic Res.* **2**, 93-100.

25 **Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A. and Lukyanov, A.S.** (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucl. Acids Res.* **23**, 1087-1088.

30 **Slatter, R.E., Dupree, P. and Gray, J.C.** (1991) A scaffold-associated DNA region is located downstream of the pea plastocyanin gene. *Plant Cell* **3**, 1239-1250.

35 **Southern, E.** (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503.

- Southern, E. (1979) Gel electrophoresis of restriction fragments. *Methods of Enzymol.* **68**, 152.
- 5 Spiker, S. and Thompson, W.F. (1996) Nuclear matrix attachment regions and transgene expression in plants. *Plant Physiol.* **110**, 15-21.
- Stein, G.S., Lian, J.B., Dworetzky, S.I., Owen, T.A., Bortell, R., Bidwell, J.P. and van Wijnen, A.J. (1991) Regulation of transcription-factor activity during growth and differentiation: Involvement of the nuclear matrix in concentration and localization of promoter binding proteins. *J. Cell. Biochem.* **47**, 300-305.
- 15 Stief, A., Winter, D.M., Stratling, W.H. and Sippel, A.E. (1989) A nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature* **341**, 343-345.
- 20 Travers, A.A. (1990) Why bend DNA. *Cell* **60**, 177-180.
- Van der Geest, A.H.M., Hall, G.E., Jr., Spiker, S. and Hall, T.C. (1994) The b-phaseolin gene is flanked by matrix attachment regions. *Plant J.* **6**, 413-423.
- 25 Van der Geest, A.H.M. and Petolino, J.F. (1998) Expression of a modified green fluorescent protein gene in transgenic maize plants and progeny. *Plant Cell Rep.* **17**, 760-764.
- 30 von Kries, J.P., Buhrmester, H. and Stratling, W.H. (1991) A matrix/scaffold attachment region binding protein; Identification, purification and mode of binding. *Cell* **64**, 123-135.

Wolffe, A.P. (1994) The transcription of chromatin templates.
Curr Opinion Gen Dev **4**, 245-254.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985)

- 5 Improved M13 phage cloning vectors and host strains:
Nucleotide sequences of the M13mp18 and pUC 19 vectors. *Gene*
33, 103-119.

Claims

1. An isolated DNA molecule comprising bp 11 to 309 of SEQ ID NO: 1.

5

2. A DNA construct comprising, in the 5' to 3' direction: a transcription initiation region functional in plant cells, a structural gene operatively associated with the transcription initiation region, a 3' untranslated region, and a matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 positioned either 5' to said transcription initiation region or 3' to said structural gene.

10

3. A DNA construct of claim 2 wherein said matrix attachment region comprises two or more tandem copies of bp 11 to 309 of SEQ ID NO:1.

15

4. A DNA construct of claim 2 wherein a first matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 is positioned 5' to said transcription initiation region and a second matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 is positioned 3' to said structural gene.

20

5. A DNA construct of claim 4 wherein each of said matrix attachment regions is comprised of two or more tandem copies of bp 11 to 309 of SEQ ID NO:1.

25

6. A method of making recombinant plant cells having increased expression of structural genes introduced therein which comprises:

30 transforming a plant cell capable of regeneration with a DNA construct of claim 2.

35

7. A method of claim 6 wherein the matrix attachment region in said DNA construct comprises two or more tandem copies of bp 11 to 309 of SEQ ID NO:1.

8. A method of claim 6 wherein the DNA construct includes a first matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 positioned 5' to said transcription initiation region and a second matrix attachment region comprised of bp 11 to 309 of SEQ ID NO: 1 positioned 3' to said structural gene.

5
9. A method of claim 8 wherein said first and second matrix attachment regions are each comprised of two or more tandem copies of bp 11 to 309 of SEQ ID NO:1.

10
10. A transformed plant cell containing DNA of bp 11 to 309 of SEQ ID NO:1.

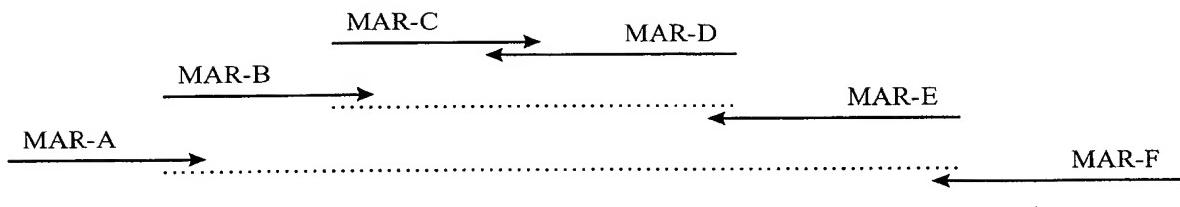


FIG 1

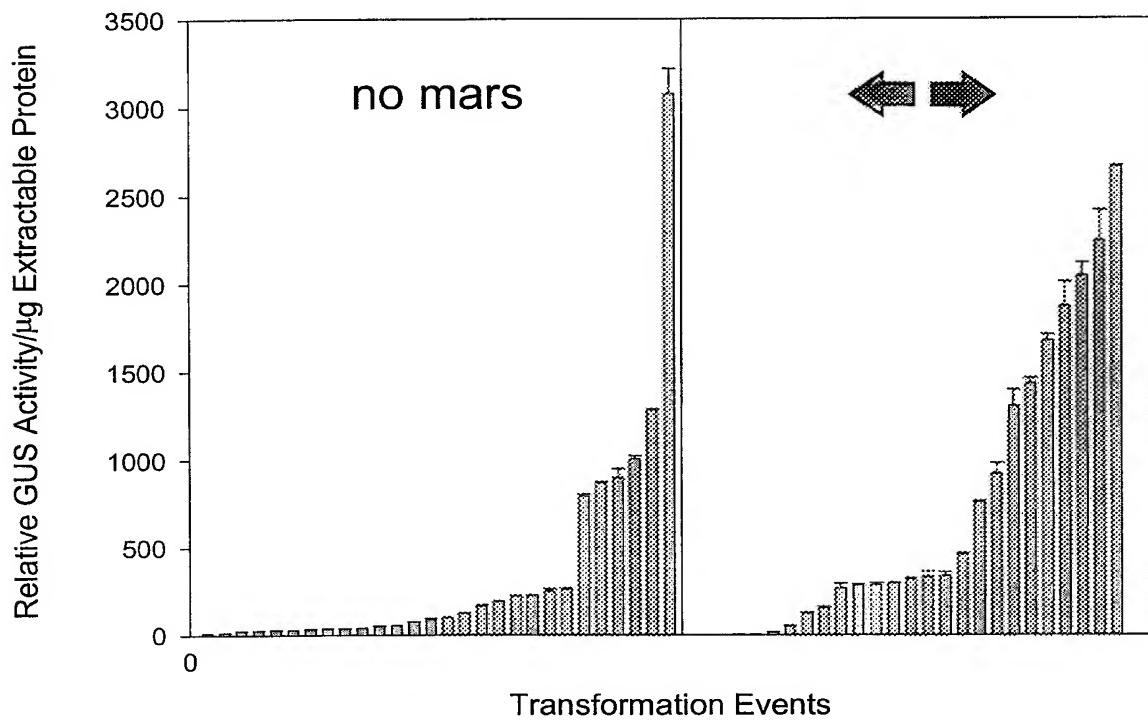


FIG 3

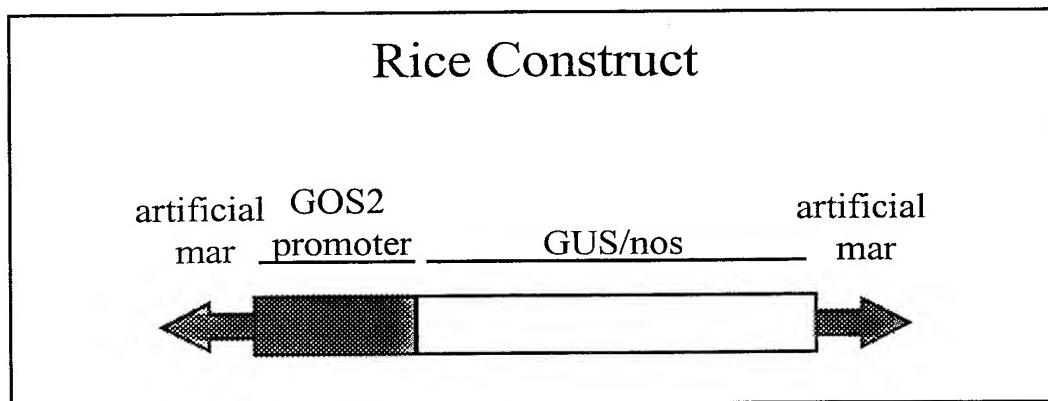


FIG 2

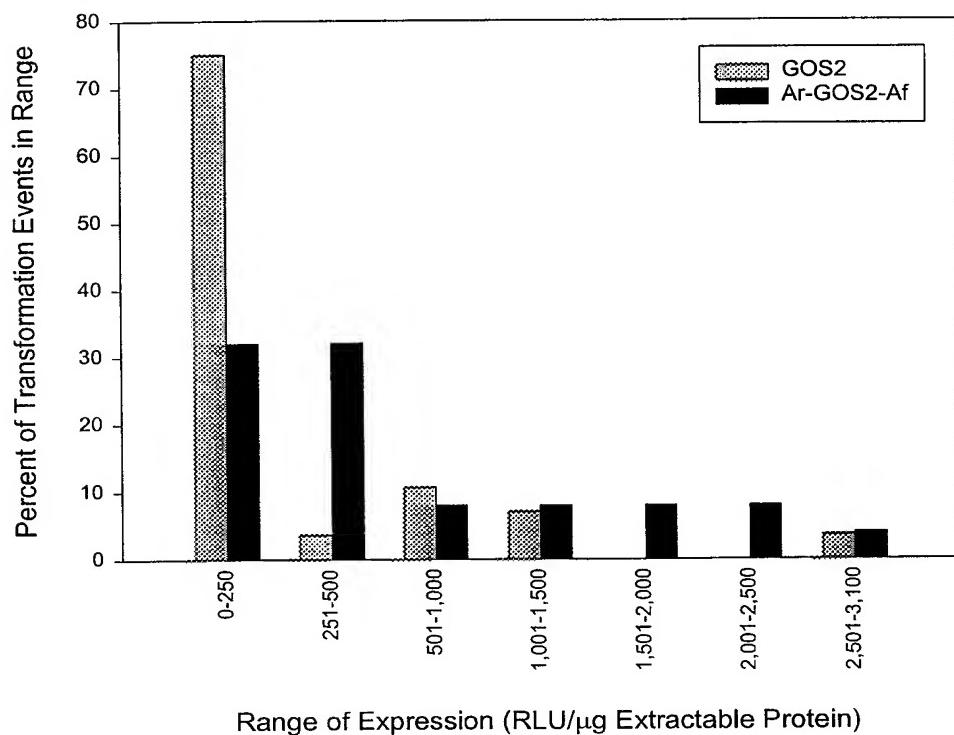


FIG 4

2 / 4

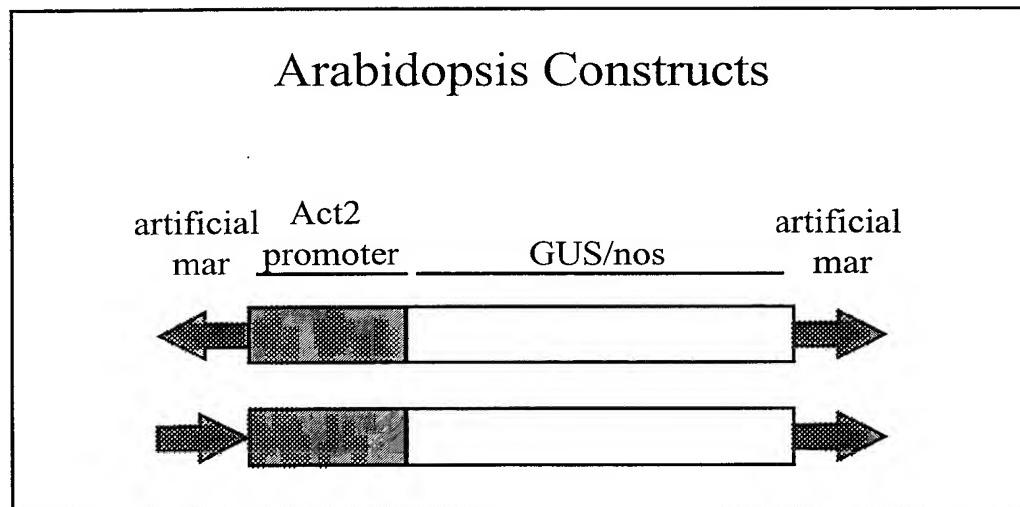


FIG 5

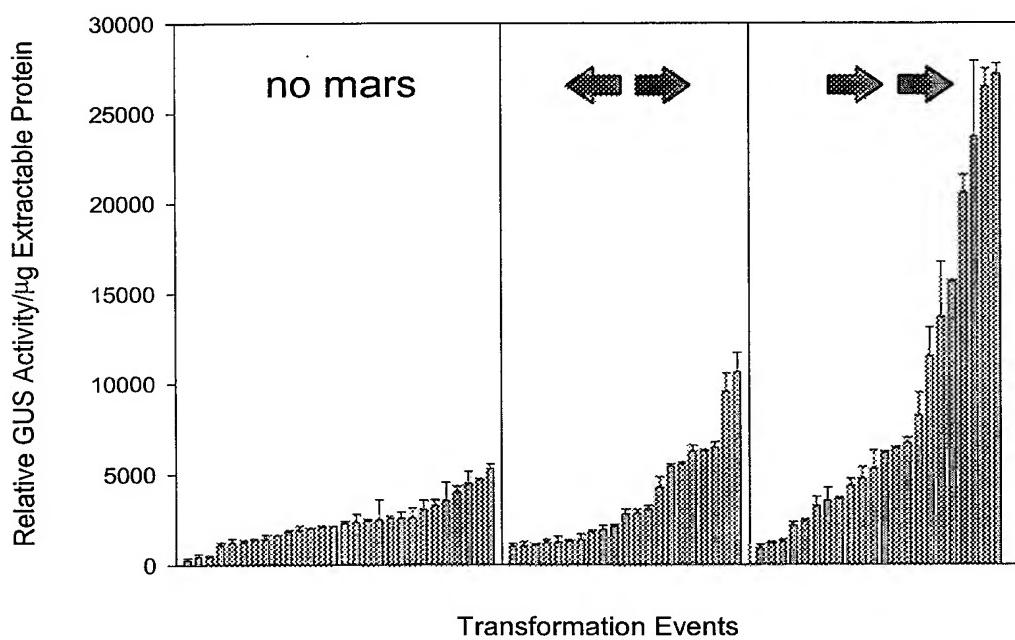


FIG 6

3 / 4

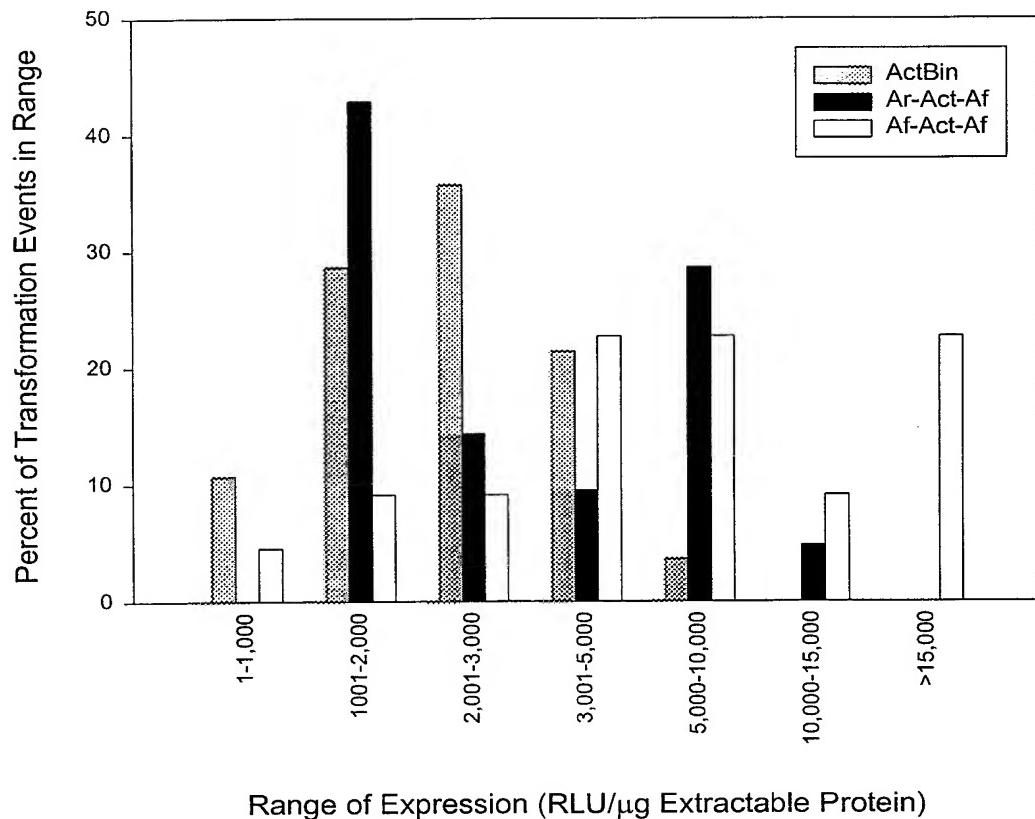


FIG 7

4 / 4

SEQUENCE LISTING

<110> van der Geest, Apolonia H. M.
Ainley, W. Michael
Cowen, Neil W.
Welter, Mary E.
Woosley, Aaron T

<120> ARTIFICIAL MATRIX ATTACHMENT REGION FOR INCREASING
EXPRESSION OF GENES INTRODUCED IN PLANT CELLS

<130> 50545

<140>
<141>

<150> US 60/110,437
<151> 1998-12-01

<160> 33

<170> PatentIn Ver. 2.0

<210> 1
<211> 327
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial MAR

<400> 1
ttccagatct cgatatctt aatttcta atat tagaa ggggtt taga tttatata tg 60
acgtccatga aaaaaa attt taaa acgata ggccagctcc aaagaatata ttccctt gga 120
gctggtaat attaattt agt ccttc cttt ctaatata ttgcata ttgattcct 180
ttttaagatt atat agctcc atgcca agt gacttc taa tatattttat gaagt cagca 240
aaat tttaa aaggcacact tgactt ggta tatatttata aatgtt taaa cttaatttca 300
cttgtaaaac tcttgcagga tccgtgc 327

<210> 2
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ARBP site

<400> 2
at ttc a stt tg taaa 15
at ttc a stt tg taaa

<210> 3
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ARBP site

<400> 3
tgcagg tgc ctt 13
tgcagg tgc ctt

<210> 4
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ARBP site

<400> 4
tgggggtgta aaa 13

<210> 5
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ATF site

<400> 5
tgacgtccat g 11

<210> 6
<211> 5
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: BEAF-32 site

<400> 6
cgata 5

<210> 7
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: topoisomerase II site

<400> 7
cnngyngktn yny 13

<210> 8
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: topoisomerase II site

<400> 8
asmatgcgyw yatcrt 16

<210> 9
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: topoisomerase
II site

<400> 9
gtnwakattn atnnr 15

<210> 10
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: unwinding
sequence

<400> 10
aatatattt 9

<210> 11
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 11
ttctaatata t 11

<210> 12
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 12
ataatcttc 9

<210> 13
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 13
ttattattta 10

<210> 14
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 14
tataaaaa 8

<210> 15
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 15
aagattatat a 11

<210> 16
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 16
ttttaatgag ataataa 17

<210> 17
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SATB1 site

<400> 17
tataatcttc 10

<210> 18
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: bending DNA site

<400> 18
aaannnnnnn aaa 13

<210> 19
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: exemplary oligo A/T tract

<400> 19
aaaaaaaa 7

<210> 20

<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-A
oligonucleotide

<400> 20
ttccagatct cgatatcttt aatttctaat atattttagaa ggggtttaga ttta 54

<210> 21
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-B
oligonucleotide

<400> 21
gaagggggttt agatttatat atgacgtcca tgaaaaaaaaa tttaaaacg ataggccagc 60
tcc 63

<210> 22
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-C
oligonucleotide

<400> 22
cgataggcca gctccaaaga atatattcc ctggagctgg taaatattaa ttagtcctct 60
ccc 63

<210> 23
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-D
oligonucleotide

<400> 23
cttggcatgg agctataataa tctaaaaag gaatcaaaaa tatcgaaaaa tatattagaa 60
gggagaggac taatt 75

<210> 24
<211> 74
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-E
oligonucleotide

<400> 24
ccaaagtcaag tgtgctttt aaaaattttg ctgacttcat aaaatatatt aggaagtcaag 60

cttggcatgg agct 74

<210> 25
<211> 74
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: MAR-F
oligonucleotide

<400> 25
gcacgatcc tgcaagagt ttacaagtga aattaagttt aaacatttat aaatatatac 60
caagtcagt gtgc 74

<210> 26
<211> 668
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 3' MAR in
ArActAf

<400> 26
gggggatctc gatatctta atttctaata tatttagaag gggtttagat ttatatatga 60
cgtccatgaa aaaaaattt aaaacgatag gccagctcca aagaatataat ttccctggag 120
ctggtaataa ttaattagtc ctctcccttc taatataattt ttcgatattt ttgatccctt 180
tttaagatta tatagctcca tgccaagctg acttccta atatttatg aagtcagcaa 240
aattttaaa aagcacactt gacttggtat atatttataa atgtttaaac ttaatttcac 300
ttgtaaaact cttgcaggat ctcgatatct ttaatttcta atatatttag aagggttta 360
gatttatata tgacgtccat gaaaaaaaaat tttaaaacga taggcccagct ccaaagaata 420
tatttcctg gagctggtaa atattaatta gtcctctccc ttctaatata ttttcgata 480
tttttatttc ctttttaaga ttatatagtt ccatgccaag ctgacttcctt aatatattt 540
atgaagtcag caaaattttt aaaaagcaca cttgacttgg tatataattt taaatgttta 600
aacttaattt cacttgtaaa actcttgcag gatccacttag ttctagagcg gccgccaccg 660
cggtgag 668

<210> 27
<211> 9361
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:pGOS2-hpt

<400> 27
gshttcgcgc gtttcggta tgacggtaa aacctctgac acatgcagct cccggagacg 60
gtcacagctt gtctgtaaac ggatgcccgg agcagacaag cccgtcaggg cgcgtcagcg 120
ggtgtggcg ggtgtcgggg ctggcttaac tatgcgcac cagagcagat tgtaactgaga 180
gtgcaccata tgcgggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcattcagg 240
cgccattcgc cattcaggct gcgcactgt tgggaagggc gatcggcgc ggccttccg 300
ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagtgg gtaacgcca 360
gggtttccc agtcacgcg ttgtaaaacg acggccagtg aattcccgat cgatctgatg 420
acatacatgtt caccgcgcgc gataattttt cctagttgc gcgcstatatt ttgtttctt 480
tcgcgtatta aatgtataat tgcgggactc taatcataaa aacccatctc ataaataacg 540
tcatgcatta catgttaattt attacatgtt taacgttaattt caacagaaat tataatgtt 600
tcatgcgaag accggcaaca ggattcaatc ttaagaaact ttattgccaa atgttgaac 660
gatcggggaa attcgagctc ttactattcc tttgcctcg gacgagtgtctt ggggcgtcgg 720
tttccactat cggcgagttac ttctacacag ccatcgtcc agacggccgc gcttctgcgg 780
cgatgttgc tacgccccgac agtcccgctt ccggatcgga cgattgcgtc gcatcgacc 840

tgcgccaag ctgcatacatc gaaattgccg tcaaccaagc tctgatagag ttggtcaaga 900
 ccaatgcgga gcataatacgcc cgaggagccgc ggcgatcctg caagctccgg atgcctccgc 960
 tcgaagttagc gcgtctgctg ctccatacaa gccaaccacg gcctccagaa gaagatgttg 1020
 ggcacctcgat attggaaatc cccgaacatc gcctcgctcc agtcaatgac cgctgttatg 1080
 cggcattgt cggtcaggac attgttggag cggaaatccg cgtgcacgag gtgcccggact 1140
 tcgggcagt ctcggccca aagcatcagc tcatcgagag cctgcgcgac ggacgcactg 1200
 acggtgtcgt ccatcacagt ttgccagtga tacacatggg gatcagcaat cgcgcatatg 1260
 aaatcacgcc atgttagtga ttgaccgatt cttgcggtc cgaatgggcc gaaccgcctc 1320
 gtctgctaa gatcgccgc agcgatcgca tccatggcct ccgcgaccgg ctgcagaaca 1380
 gcccggcgtt cggttccagg caggcttgc aacgtgacac cctgtgcacg gcgggagatg 1440
 caataggtca ggctctcgct gaattccccaa atgtcaagca cticccgaat cgggagcgcg 1500
 gccgatgcaa agtgcgcata aacataacga tctttaga aaccatcgcc gcagcttatt 1560
 acccgccagga catatccacg ccctcctaca tcgaagctga aagcacgaga ttcttcgccc 1620
 tccgagagct gcatcaggcggc ggagacgctg tcgaactttt cgatcagaaa cttctcgaca 1680
 gacgtcgcgg tgagttcagg cttttcata tctcacgaga tctgtatccc gtgcctggaa 1740
 caaatggccc cagatccgtc cgcaagtcga cgggtccagg aaagcaatcg catagtcaag 1800
 ctaaatcatc aagatgcaaa ctttcgccc ttgctaaaca cgtaaaatt cgaatggaca 1860
 tgtgtggagc agcaaaggag ctttccccaa aattactcaa cgaatcataa accaagatta 1920
 gtcagatcaa gagacagagg agaaacaagg cggacccttgcacttgcgtt gggattgcc 1980
 ctgacttgggt ggtgctggta tattagggat agggttgcctc statccacag ctgttccacc 2040
 aaatatcagc tcctccgtgg actgcctgt cgagccttca gctggatcct ctagagtccc 2100
 ccgtttctc tccaaatgaa atgaacttcc ttatataagag gaagggtctt gcgaaggata 2160
 gtgggattgt ggcgtatccc ttacgtcagt ggagatatac catcaatcca cttgccttga 2220
 agacgtgggt ggaacgttcc cttttccac gatgctcctc gtgggtgggg gtccatctt 2280
 gggaccactg tcggcagagg catcttcaac gatggcctt ctttatcgc aatgatggca 2340
 tttttaggag ccaccccttcttccactat cttcacaata aagtgcacaga tagctggca 2400
 atggaatccg aggagggttc cggatattac ctttgccttga agacgtgggt ggaacgttcc 2460
 catcaatcca ttttgccttga agacgtgggt gtcacgttcc gatgctcctc 2520
 gtgggatgggg gtccatctt gggaccactg tcggcagagg catcttcaac gatggcctt 2580
 ctttatcgc aatgatggca tttttaggag ccaccccttcttccactat cttcacaata 2640
 aagtgcacaga tagctggca atggaatccg aaggagggttc cggatattac ctttgccttga 2700
 aaagtctcca cccatgcaga tctgcaggca tgcaagctgc ggcgcgtt aaacaattcc 2760
 cgatcgatct agtaacatag atgacaccgc ggcgcataat ttatcttgcgtt ttgcgcgtt 2820
 tattttgttt totatcgctt attaaatgtt taattgcggg actctaatac taaaaaccca 2880
 tctcataaat aacgtcatgc attacatgtt aattattaca tgcttaacgt aattcaacag 2940
 aaattatatg ataatcatcg caagaccggc aacaggattc aatcttaaga aactttattg 3000
 ccaaatgtttt gaacgatcg gggaaattcgaa gctctccaat tccccaccga ggctgttagcc 3060
 gacgatgggtt cggcaggaga gttgttgcatttgccttccactatccacttccatgc 3120
 ccgaagttca tgccagtccaa ggcgttttgc accttctctg ccgtttccaa atgcgcgtt 3180
 tcgcgagtga agatccctt cttgttaccg aacaggattc aatcttaaga aactttattg 3240
 aaatccggca aattccatc ctgttccaccg acgacgcgc tgacgcgttccatgc 3300
 tgatcatat ccagccatgc acactgatac tcttcacttccatgc 3360
 tgcagcccggtt ctaacgtatc cacgcccgtat tcggtgatga taatcggttgc 3420
 tcctgccagg ccagaaggttc ttttccactt accttctctg ccgtttccaa atgcgcgtt 3480
 tggacatacc atccgtataa acggttcagg cacagcacat caaagagatc gctgtatggta 3540
 tcggtgttgcg cgtgcgcgaa cattacatttgc acgcaggatgc tcggacgcgtt cgggtcgagt 3600
 ttacgcgttgc ttccgccttgc tggcggaaata tttccgttgc ctgcggacgc ggtatccgg 3660
 tcgttggcaat tactccacat caccacgcgtt ggggtgggtt tgcacgcgc tatcagctt 3720
 ttaatcgctt gtaagtgcgc ttgttgcgtt tccccgttgc ctgccttc gctgtacagt 3780
 tctttccgttgc ttgttgccttgc ttcgaaacca atgcctaaag agaggtaaa gccgacagca 3840
 gcagtttcat caatcaccac gatgccatgt tcatcttgc acgtcgatccatgc 3900
 taagggttaat gcgaggatcgc gtagggatgtt gccccaaatcc agtccattaa tgcgttgcgtt 3960
 tgcaccatca gcacgttatac gaatcccttgc ccacgttaatc cgcgcatttc atgacgcacca 4020
 aagccagtaa agtagaacgg tttgtggta atcaggaaact gtcgccttgc cactgcctt 4080
 gaccggatgc cgacgcgcgaa cgggttagata tcacactctg tctggctttt ggctgtgcgc 4140
 cacaggatcat agagataacc ttccaccgcgtt tgccagaggatcgc gggatttccac cacttgc 4200
 gtcccgcttag tgccttgccttgc agtgcacacc acctgttgc tccgcacatcgc cagttcaacgc 4260
 ctgacatcac cattggccac cacctgcgcgac tcaacagacgc cgtggttaca gtcttgcgc 4320
 acatcggtca ccacgggtatcgttgc atcgttccacc cagggttgc gctgttgcgtt gaggattacgc 4380
 ctgcgttgc ttccggcata gttaaagaaa tcatggaaatc agactgttccatgc 4440
 ttttcgttgcg taatcaccat tccggccggg atagtcgttgc agttcgttgc gttgttgcaca 4500

caaacggta tacgtacact tttccggca ataacatacg gcgtgacatc ggcttcaa 4560
 ggcgtatagc cgccctgatg ctccatcaact tcctgattat tgacccacac tttggcgtaa 4620
 tgagtgaccg catcgaaacg cagcacgata cgctggctg cccaacctt cggataaaag 4680
 acttcgcgct gataccagac gttgccgca taattacgaa tatctgcattc ggcgaactga 4740
 tcgttaaaac tgcctggcac agcaattgcc cggcttctt gtaacgcgct ttcccaccaa 4800
 cgctgatcaa ttccacagtt ttcgcgatcc agactaatg cccacaggcc gtcgagttt 4860
 ttgatttcac gggttgggt ttctacagga cggaccatgg aactttgctg gtgaaagtgg 4920
 caagctacac caagaactgc acaaaggaaa tgaatcataa gaacaaggca tataactatc 4980
 ccgattacag cttcataaaat ttcttctgt aatcaagcag tcaaggaata accaaaaaga 5040
 aacttctaca ggtagatatac agaggataat cgatagataa tcgatgtgaa tttgaaaaca 5100
 aaatttgagga cagttcatgc cagacttca gctcagaata atgaaggggat gagagcta 5160
 aaaaaaaaaata atccaaatga atactgctta tgctacagtt catttcataat aattaaaaat 5220
 ggagatcaga aatcgatataa gttcttctcc tgactaaact ataggggtat tacattaa 5280
 ctgaactaca gctaggataa cgctataaaa gcattatttgc tagcaatcaa ttcatgtAAC 5340
 tgaaccagaa agtgacttt taagatattt gggaaaaaaa attctgggac taaagcaa 5400
 cggaaagaaca gggatcccc tggctctgag gattataact gttccatga atttcgttat 5460
 ggggactact gtatctaaac aagccagctg cgaaattttgc gacaggctt agaataatc 5520
 attcaatctc atcaacggga cggcttcaaa agttggatta ttttggatca ataggaaata 5580
 aacaaaggat agttcgtca aatcgagaag catcaactacc agaatacgagg accaaacaac 5640
 cgtactttta ttacaccaag caaaatcacc ggtgctctga ttttacctca aacaaaggat 5700
 actcacaaaa tcgcaagattt cccgtaccctt aaccattca tttccataga gctctccaga 5760
 cgattgaaaa ccataactact aatcaaaaccg aaccgataac atgcaacatc aagaacccct 5820
 ctatccaaa tccaagaaca ggaatcatca cagatacaga tccccttcc taacatcaac 5880
 gcccgtacta cacaaccttag aacaataaaat ccaagaacaa ccgaaggggca cataccctgt 5940
 gaggaggagg tggaggaggaa aagagatcga ccaagaactc gaccggaaaga tatggatcga 6000
 agaaggcggt cgctcggtt ctgctagtcg cgtgtccttgc gtgtctccc tcttcttggc 6060
 ttggatgcct cctatataga gggggaaaaa ggggggagga atttatagat gggaggagga 6120
 ggatgcttgg ttttctttgc ctctccttctt ctctccttgc cggaaaggct gctgttaaaa 6180
 ggttggcg gacttgcgtt cctccgttag atcatcgattt tttgtgggtt gttctgtggg 6240
 cagccactct gttgttgcct gtgtgccc aatgggagat tgccgtcgc gcacgagcaa 6300
 aattctttt tttttgtat gtggaaaaacc taaatagttc gtttcatact tttcatgcta 6360
 ttctgtaaaa ttattttttgc tatttttagat attattaaaaat gtggctgtgt gatggtgag 6420
 tgcttgaattt cagatatttc tacctaaatg tattaaatag gcgagtgata ttagagatgt 6480
 gttgttagtt gaacgtgtat tgaggaggta cactcacaca tgacatgag cacaaggatgt 6540
 gtgcgttaatc acctgcattc taatcgaaaaa aagataaata ataaaaataa gtcaattgtc 6600
 tttaattactt aaataaaaaat tggatggag taagacatgt ccttaatgat gtgcgtatgt 6660
 acgaatgcac aaactataaa attatataat tataatgttta aatctttgccc aatacgttca 6720
 gaatatcttc attctattttt tttttttttttt atctctcttctt ttacccattt agttcagcta 6780
 gaaagattttt tttttcaag aagagtttgc tgacaattat ggctacctcg aataatttaa 6840
 cttcatgaca gagatgtgaa cgtatattct aagcaataat gatttcattt tcccacttaa 6900
 ttactaagga aaacgaaact agttagcgat ctttttttta tttagcccgaa ttgcccactaa 6960
 agtaggtgga tgagtttttc ttgcataatgtt ctatccatca ggcgtacata tataaggct 7020
 cattttatattt ttagcacacg ttccttatat tgtagttatgggtggaaaaa cgggtcagaa 7080
 actgcaggca tgcaagtttcc cccggaaatgc ggccggcgc cgcagctggc gtaatcatgg 7140
 tcatacgctgt ttccgtgttgc aaattgttat ccgctcacaa ttccacaccaa catacgagcc 7200
 ggaagcataa agtgtaaacgc ctgggggtgccc taatggatgtgaa gctaactcactt attaatttgcg 7260
 ttgcgtcac tggccgtttt ccagtcggga aacctgtcgat ggcagctcgat ttaatgtatc 7320
 ggc当地acgca cggggagagg cgggttgcgtt atggggcgctt cttccgttcc ctcgtcact 7380
 gactcgctgc gctcggtcgat tcggctgcgg cggcggat cagctcactt aaaggcggt 7440
 atacggttat ccacagaatc agggataac gcaaggaaaga acatgtgagc aaaaggccag 7500
 caaaaggcca ggaaccgtaa aaaggccgcg ttgtgtcgat tttccatag gctccggcccc 7560
 cctgacgagc atcacaaaaaa tcgacgctca agtcagatgtt ggcgaaaccc gacaggacta 7620
 taaagatacc aggctttcc cccctggaaatc tccctcgatgc gctctctgt tccgaccctg 7680
 ccgcttaccg gatacctgtc cgccttccgc ccttcggaa ggttggcgat ttctcaatgc 7740
 tcacgctgtt ggtatctcag ttccgtgtat gtcgttcgtt ccaagactggg ctgtgtgcac 7800
 gaaccccccgg ttcagccgaa ccgtcgccgc ttatccgttactatcgat ttagtccaa 7860
 ccggtaagac acgacttatac gccactggca gcaaggactt gtaacaggat tagcagagcg 7920
 aggtatgttag gcggtgtac agatgttttgc aagtgggtgc ctaactacgg ctacactaga 7980
 aggacagttat ttggatctgc cgctctgttgc aagccagttt ccttcggaa aagagttgg 8040
 agctcttgcattt ccggcaaaaca aaccaccgtt ggttagcggtt gttttttgtt ttgcaagcag 8100
 cagattacgc gcagaaaaaaa agatctcaa gaagatcctt tgatctttc tacggggct 8160

gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 8220
 atcttacact agatccttt aaattaaaaa tgaagttta aatcaatcta aagtatatat 8280
 gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 8340
 tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 8400
 gagggcttac catctggccc cagtctgca atgataaccgc gagaccacccg ctcaccggct 8460
 ccagatttat cagcaataaa ccagccagcc ggaaggggccg agcgcagaag tggtcctgca 8520
 actttatccg cctccatcca gtctattaat tggtgccggg aagcttagat aagtagttcg 8580
 ccagtaataa gtttgcgcaa cggttgttgc attgctacag gcacgtggg gtcacgctcg 8640
 tcggttgta tggcttcatt cagctccggc tcccaacgat caaggcgagt tacatgatcc 8700
 cccatgttgt gcaaaaaaagc ggttagctcc ttccgtcctc cgatcgttgt cagaagtaag 8760
 ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 8820
 ccatccgtaa gatgctttc tgtgactggg gagactcaa ccaagtcatt ctgagaatag 8880
 tgtatgcggc gaccgagttg ctctgcccgc gcgtcaatac ggataatac cgcgcacat 8940
 agcagaacct taaaagtgt catcattggaa aacgttctt cggggcgaaa actctcaagg 9000
 atcttaccgc tggttagatc cagttcgatg taacccactc gtgcacccaa ctgatctca 9060
 gcatcttta ctttcaccag cggttctggg tgagcaaaaa caggaaggca aaatgccgca 9120
 aaaaaggaa taagggcgac acggaaatgt tgaataactca tactcttcct tttcaatat 9180
 tattgaagca tttatcaggg ttattgtctc atgagccgt acatatttga atgtatttag 9240
 aaaaataaac aaatagggtt tccgcgcaca ttcccccga aagtgcacc tgacgtctaa 9300
 gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gcccttcgt 9360
 c 9361

<210> 28

<211> 10629

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:pArGOS2Af-hpt

<400> 28

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
 cagctgtct gtaagcggat gcccggagca gacaagcccg tcagggcgcg tcagcggtg 120
 ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgtt ctgagagtgc 180
 accatatgcg gtgtgaaata cccgacagat gcgtaaaggag aaaataccgc atcaggcgcc 240
 attcgccatt caggctgcgc aactgttggg aaggggcgatc ggtgcgggccc tttcgctat 300
 tacgcccggc ggcgaaaggg ggtatgtgtca aaggcgatt aatgtgggtt acgcccagggt 360
 tttcccgatc acgacgttgtt aaaaacgacgg ccagtgaatt cccgatcgat cttagtaacat 420
 agatgacacc ggcgcgcata atttaccta gtttgcgcgc tatattttgt tttctatcgc 480
 gtattaaatg tataattgcg ggactctaat cataaaaaacc catctctataa ataacgtcat 540
 gcattacatg ttaattattt catgcttaac gtaattcaac agaaattata tgataatcat 600
 cgcaagaccg gcaacaggat tcaatcttaa gaaatttat tgccaaatgt ttgaacgatc 660
 gggaaattc gagctcttac tattccttg ccctcgacg agtgcgtggg cgtcggtttc 720
 cactatcgcc gtagtacttc acacagccat cggccagac gggcgccgtt ctgcgggccc 780
 tttgtgtacg cccgacagtc ccggctccgg atcggacgt tgctgcgcat cgaccctgcg 840
 cccaaactgc atcatcgaaa ttggcgtaa ccaagctctg atagagtttgc tcaagaccaa 900
 tgcggagcat atacgccccg agccgcggcg atcctgcaag ctccggatgc ctccgctcga 960
 agtagcgcgt ctgctgctcc atacaagcca accacggcct ccagaagaag atgttggcga 1020
 cctcgattt ggaatccccg aacatcgccct cgctccagtc aatgaccgcgt gttatgcggc 1080
 cattgtccgt caggacattt ttggagccga aatccgcgtg caccggatgc cggacttcgg 1140
 ggcagtccctc ggcggaaagc atcagctcat cgagaggctg cgcgcacggac gcactgacgg 1200
 tgtcgccat cacagtttc cagtgataca catggggatc agcaatcgcc catatgaaat 1260
 cacgcacatgt agtgtatttttgc ccgattccctt gccgtccgaa tggggccgaac ccgctcgtct 1320
 ggctaagatc ggcggcagcg atcgcacatcca tggcccccgc gaccggctgc agaacagcgg 1380
 gcagttcggt ttcaggcgagg tcttgcaacg tgacaccctg tgacacggcg gggatgcatt 1440
 aggtcaggct ctcgctgaat tcccaatgt caagcacttc cggaaatcgcc agcgcggccg 1500
 atgcaaaatgt ccgataaaaca taacgatctt tgtagaaacc atcggcgcaag ctatttaccc 1560
 gcaggacata tccacgcctt cctacatcgaa agctgaaagc acgagattct tcgcctccg 1620
 agagctgcgt caggtcgagg acgctgtcgaa acttttcgt cagaaacttc tcgcacagacg 1680
 tcgcgttgag ttcaggctt ttcatatctc acgagatctg tatcccgatgc ctggaaacaaa 1740
 tggcccccaga tccgtccgca gtcgcacggg tccagggaaag caatcgatca gtcaagctaa 1800

atcatcaaga tgcaaactt tcgcccgtc taaacacggt aaaattcgaa tggacatgtg 1860
 tggagcagca aaggagctt cccaaaatt actcaacgaa tcataaacca agattagtca 1920
 gatcaagaga cagaggagaa acaaggcggc ccttcgcact ttagtctgggg attgccctga 1980
 ctgggtggtg ctggtatatt agggataggg ttgctctat ccacagcttgc tccaccaa 2040
 atcagtcctt ccgtggactg ccttgcgg ccttcagctg gatcctctag agtccccgt 2100
 gttctctcca aatgaaatga acttccttata atagaggaag ggtcttgcga aggatagtgg 2160
 gattgtgcgt catcccttac gtcagtggag atatcacatc aatccacttgc ctttgaagac 2220
 gtgggtggaa cgtcttctt ttccacgatc ctcctcgtgg gtgggggtcc atcttggga 2280
 ccactgtcgg cagaggcatc ttcaacgatc gccttcctt tatcgcataatg atggcatttgc 2340
 taggagccac ctccctttc cactatcttca acaataaagt gacagatagc tggcaatgg 2400
 aatccgagga ggttccggta tattaccctt tggtaaaaag tctccatcgatc tgatcacatc 2460
 aatccacttgc ttgtgaagac gtgggtggaa cgtcttctt ttccacgatc ctcctcgtgg 2520
 gtgggggtcc atcttggga ccactgtcgg cagaggcatc ttcaacgatc gccttcctt 2580
 tatcgcataatg atggcatttgc taggagccac ctccctttc cactatcttca acaataaagt 2640
 gacagatagc tggcaatgg aatccgagga ggttccggta tattaccctt tggtaaaaag 2700
 tctccaccca tgcagatctg cagcatcgatc agctgcggcc gctctagaac tagtggatcc 2760
 tgcaagagtt ttacaagtga aattaagtt aaacattat aataatatac caagtcaatg 2820
 gtgctttta aaaatttgc tgacttcata aataatatta ggaagtgc tggcatgg 2880
 gctatataat cttaaaaagg aatcaaaaat atcgaaaaat atattagaag ggagaggact 2940
 aattaatatt taccagctcc agggaaatatt atctttggta gctggctat cgttttaaaa 3000
 tttttttca tggacgtcat atataaatct aaacccttc taatataat agaaattaaaa 3060
 gatatcgaga tcctgcaaga gtttacaag tgaattaaatg tttaaacatt tataaatata 3120
 taccaagtca agtgtgcattt taaaaattt tgctgacttc ataaaatata ttaggaagtc 3180
 agcttggcat ggagctataat aatcttaaaa aggaatcaaa aatatcgaaa aatataatag 3240
 aaggagagg actaattaat attaccagc tccagggaaa tatattctt ggagctggcc 3300
 tatcgttta aaatttttt tcattggacgt catatataaa tctaaacccc ttctaaatat 3360
 attagaaatt aaagatatcg agatccccca aacaattccc gatcgatcta gtaacataga 3420
 tgacaccgcg cgcgataatt tattcttagtt tgcgctat attttgtttt ctatcgctg 3480
 ttaaatgtat aattgcggta ctctaattcat aaaaacccat ctcataaata acgtcatgca 3540
 ttacatgtta attattacat gcttaacgta attcaacaga aattatata taaatcatg 3600
 aagaccggca acaggatca atcttaagaa actttattgc caaatgttt aacgatcg 3660
 gaaattcggag ctctccaatt cccaccggag gctgtagccg acgtggatgc gccaggagag 3720
 ttgttgattt attgttgc tccctgctgc gtttttcac cgaagttcat gccagtc 3780
 cggttttgc gcaaaaaagc cgccgacttc ggttgcggt cgcgagtggaa gatcccttc 3840
 ttgttaccgc caacgcgcaaa tatgccttgc gaggtcgcaaa aatcgccgaa attccatacc 3900
 tgttaccgc cgacggcgct gacgcgatca aagacgcggt gatacatatc cagccatgca 3960
 cactgatact cttcacttca catgtcggtg tacattgagt gcagccggc taacgtatcc 4020
 acgccgtatt cggtgatgat aatcggtca tgcagtttct cctgcccaggc cagaagttct 4080
 tttccagta cttctctgc cgtttccaaa tcgcccgtt ggacatacca tccgtaataa 4140
 cggttcaggc acagcacatc aaagagatcg ctgatggat cggtgtgagc gtgcgagaac 4200
 attacattga cgcaggtgat cggacgcgatc gggtcgagtt tacgcgttgc ttccgccc 4260
 ggcgaaatat tcccgtgcac ttgcggacgg gtagccgtt cggtggcaat actccacatc 4320
 accacgcttgc ggtggttttt gtcacgcgatc atcagcttt taatcgccctg taagtgcgct 4380
 tgctgagttt ccccggttgc tgcctcttgc ctgtacagtt ctggccgtt gttggccgct 4440
 tcgaaaccaa tgcctaaaga gaggttaaaag ccgacacgcg cagtttcatc aatcaccacg 4500
 atgcacatgtt catctgccc gtcgacatc tcttcacgt aagggtatgc cgaggtacgg 4560
 taggagttgg ccccaatcca gtcattaaat gctgtgtcgatc gacccatcg cacgttatcg 4620
 aatcccttgc cacgtaaatc cgcacatctca tgacgaccaaa agccagtaaa gtagaacgg 4680
 ttgtggttaa tcaggaactg ttgccttgc actggactg accggatgcc gacgcgaa 4740
 gggtagatat cacactctgt ctggcttttgc gctgtgcgc acagttcatc gagataaccc 4800
 tcacccgggtt gccagagggtg cgatttccatc acttgcaccc tcccgctat gcctgtcc 4860
 gttgcaaccca cctgttgc tgcacatcg acgttcaacgc tgacatcacc attggccacc 4920
 acctggccagt caacagacgc gtggttacag tcttgcgcgatc catgcgttgc cacgggtata 4980
 tcgtccaccc aggtgtcggtt cggtgttagt agcattacgc tgcgtatggat tccggcatag 5040
 ttaaagaaat catggaaatgc agactgctt ttcttgcgtt ttgcgtcgatc aatcaccatc 5100
 cccggcggga tagtctgcca gttcagttgc ttgttccacac aaacgggtat acgtacactt 5160
 ttcccgccaa taacatacgg cgtgacatcg gttcaatgc gctgtatagcc gccctgtatgc 5220
 tccatcactt cctgattatt gacccacact ttgccttaat ggtgaccgc atcgaaacgc 5280
 agcacgatac gctggcttgc ccaaccttgc ggtataaaga cttcgccgtt ataccagacg 5340
 ttgcccgcatt aattacaat atctgcacatc gctgtatgc gttaaaact gcctggcaca 5400
 gcaattgccc ggcttcttgc taacgcgtt tccaccaac gctgtatcaat tccacagttt 5460

tcgcgatcca	gactgaatgc	ccacaggccg	tcgagttttt	tgatttcacg	ggttgggggtt	5520
tctacaggac	ggaccatgga	actttgctgg	tgaaaagtggc	aagctacacc	aagaactgca	5580
caaaggaaat	gaatcataag	aacaagcagt	ataactatcc	cgattacagc	ttcataaaatt	5640
tctttctgt	atcaaggcgt	caaggaataa	ccaaaaagaa	acttctacag	gtagatacaa	5700
gaggataatc	gatagataat	cgatgtgaat	ttgaaaacaa	aattgaggac	agttcatgcc	5760
agacttcag	ctcagaataa	tgaagggggt	agagctaata	aaaaaaataa	tccaaatgaa	5820
tactgcttat	gctacagttc	atttcataata	attaaaaatg	gagatcagaa	atcgataag	5880
ttcttctcct	gactaaacta	tagggattt	acctattaac	tgaactacag	ctagataac	5940
gctataaaag	cattatttgt	agcaatcaat	tcattgaact	gaaccagaaa	gtgactttt	6000
aagatattt	ggaaaaaaaa	ttctggact	aaagcaatc	ggaagaacag	ggaatcccct	6060
gttcctgagg	attataactg	tttccatgaa	tttcgtatg	gggactactg	tatctaaaca	6120
agccagctgc	gaaattttg	acaggcttaa	gaatcaatca	ttcaatctca	tcaacggac	6180
cgtctcaaa	gttggattat	ttttgttcaa	tagggataaa	acaaaggata	gcttcgtcaa	6240
atcgagaagc	atcactacca	gaatcgagga	ccaaacaacc	gtactttat	tacaccaagc	6300
aaaatcaccg	gtgctctgat	tttacctcaa	acaaaaggta	ctcacaaaat	cgcaagattc	6360
cgtaccctaa	accatttcat	ttccatagag	ctctccagac	gattgaaaac	catactacta	6420
atcaaaccga	accgataaca	tgcaacatca	agaaccctc	tatcccaa	ccaagaacag	6480
gaatcatcac	agatacagat	ccccttcct	aacatcaacg	ccctactac	acaacctaga	6540
acaataaaatc	caagaacaac	cgaaggccac	ataccctgt	aggaggaggt	ggaggaggga	6600
agagatcgac	caagaactcg	accggaagat	atggatcgaa	gaaggcggtc	gctcggcttc	6660
tgctagtcgc	gtgtccttgg	tgctctccct	cttcttggct	tgatgcctc	ctatataagag	6720
agggaaaaag	gggggaggaa	tttatacgat	ggaggaggag	gatgcttgg	tttcttgcc	6780
tctccctc	tctcctggcc	gcaaaagcctg	ctgtaaaag	gttgtgcgg	acttgctgtc	6840
ctccgttaga	tcatcgttt	ttgtgggtt	ttctgtggc	agccactctg	ttgttgccctg	6900
tgtgccaat	atgggagatt	ggcgctcg	cacgagcaa	attcttttt	tttttgtatg	6960
tgaaaaaacct	aaatagttc	tttcataactt	ttcatgtat	tctgtaaaat	tatTTTTG	7020
attttagata	ttattaaaag	tggcttgg	atggtgag	gcttgaattc	agatattgct	7080
acctaataatgt	attaaatagg	cgagtgtat	tagagatgt	ttctgtatgg	aacgtgtatt	7140
gaggaggtgc	actcacacat	gcacatgagc	acaaagtgt	tgcgtaaatg	ccttgcacat	7200
aatcgaaaaa	agataaataa	taaaaataag	tcaatttgt	ttaattacta	aataaaaatt	7260
gggatggagt	aagacatgtc	cttaatgtat	tgcgatatga	cgaatgcaca	aactataaaa	7320
ttatataatt	atatgtttaa	atcttgc	atacgttc	aatatctca	ttcttatttt	7380
ttaaaaaaaa	tctctctt	taccattga	gttcagctag	aaagatttt	ttattcaaga	7440
agagtttgat	gacaattatg	gctacctcg	ataatttaac	ttcatgtacag	agatgtgaac	7500
gtatattcta	agcaataatg	atttcatttt	cccacttaat	tactaaggaa	aacgaaacta	7560
gtgtagcgc	tctttttat	ttagccccat	tgccactaaa	gtaggtggat	gagttttct	7620
tgcatagttc	tagttatcg	cgctacat	ataaggctc	attttatatt	tagcacacgt	7680
tccctatatt	gttagttagg	gggtgaaaac	ggtgcagaaa	ctgcagcccc	ggggatctcg	7740
atatctttaa	tttctaatat	atttagaagg	ggttttagatt	tatatatgac	gtccatgaaa	7800
aaaaattttt	aaacgatagg	ccagctccaa	agaatatatt	tccctggagc	tggtaaatat	7860
taatttagtcc	tctccctt	aatatatttt	tcgatatttt	tgattcctt	ttaagattat	7920
atagctccat	gccaagctg	cttcctaata	tatTTTat	tgatgtatg	agtcgacaa	7980
agcacactt	acttggata	tatttataa	tgtttaaact	taatttact	tttttttttt	8040
ttgcaggatc	tcgatatctt	taatttctaa	tatatttata	tttttttttt	tttttttttt	8100
gacgtccatg	aaaaaaaatt	ttaaaacat	aggccagctc	caaagaatat	tttttttttt	8160
agctgtaaa	tattaattag	tcccttcct	tctaataat	tttttttttt	tttttttttt	8220
tttttaagat	tatatagttc	catgccaagc	tgacttcata	atattatata	tgaagtccac	8280
aaaattttt	aaaagcacac	ttgacttgg	atattatata	aaatgtttaa	acttaatttc	8340
acttgtaaaa	ctcttgcagg	atccactgt	tcttagacgg	ccgcgcgcgg	cagctggcgt	8400
aatcatggtc	atagctttt	cctgtgtgaa	atgtttatcc	gtcacaat	ccacacaaca	8460
tacgagccgg	aagcataaaag	tgttaaagct	gggggtgccta	atgagtgagc	taactcacat	8520
taatttgcgtt	gcgctca	cccgtttcc	agtcggaaa	cctgtgtgc	cagctgcatt	8580
aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	tggcgctct	tccgcttcct	8640
cgctcaactga	ctcgctgcgc	tcggcgtt	ggctgcggcg	agcggtatca	gctcaactaa	8700
aggcggtat	acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	8760
aaggccagca	aaaggccagg	aaccgtaaaa	aggccgcgtt	gtggcgctt	ttccataggc	8820
tccgcccccc	tgacgagcat	cacaaaatc	gacgctcaag	tcagagggtgg	cgaaacccga	8880
caggactata	aagataccag	gcgtttcccc	ctggaaagctc	cctcgtgcgc	tctccgttcc	8940
cgaccctg	gcttaccgga	tacctgtcc	cctttctccc	ttcggaaagc	gtggcgctt	9000
ctcaatgctc	acgctgttag	tatctcagtt	cggtgttagt	cgttcgtcc	aagctggct	9060
gtgtgcacga	accggccgtt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgttctt	9120

agtccaaacc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	9180
gcagagcgag	gtatgttaggc	ggtgctacag	agttcttgc	gtggtggct	aactacggct	9240
acaactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaaa	9300
gagttggtag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcggttgt	ttttttgttt	9360
gcaaggcagca	gattacgcgc	agaaaaaaaaag	gatctcaaga	agatccttg	atctttctta	9420
cggggctctga	cgttcagtgg	aacggaaaact	cacgttaagg	gattttggtc	atgagattat	9480
caaaaaggat	cttcacacctag	atcctttaa	attaaaaatg	aagttttaaa	tcaatctaaa	9540
gtatatatga	gtaaaacttgg	tctgacagt	accaatgctt	aatcagttag	gcacccatct	9600
cagcgatctg	tctatttcgt	tcatccatag	ttgcctgact	ccccgtcg	tagataacta	9660
cgatacggga	gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gaccacgct	9720
caccggctcc	agattttatca	gcaataaaacc	agccagccgg	aaggcccgag	cgcagaagt	9780
gtccctgcaac	tttatccgccc	tccatccagt	ctattaatttgc	ttgcggggaa	gttagagtaa	9840
gtagttcgcc	agttaatagt	ttgcgcaacg	ttgttgccat	tgctacaggc	atcgtgggt	9900
cacgctcgtc	gtttggatg	gcttcattca	gctccgggtc	ccaacgatca	aggcgagtt	9960
catgatcccc	catgttgtgc	aaaaaaagccgg	ttagctcctt	cggccctccg	atcgttgtca	10020
gaagtaagtt	ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	10080
ctgtcatgcc	atccgtaa	tgctttctg	tgactggtga	gtactcaacc	aagtcatct	10140
gagaatagt	tatgcggcga	ccgagttgt	cttgcggcgc	gtcaataacgg	gataataccg	10200
cggccacatag	cagaacttta	aaagtgtca	tcattggaaa	acgttcttcg	gggcgaaaac	10260
tctcaaggat	cttaccgcgt	ttgagatcca	gttcgatgt	accactcg	gcacccaaact	10320
gatcttcagc	atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	10380
atgcccaaa	aaagggaaata	agggcgcacac	ggaaatgtt	aataactata	ctcttcctt	10440
ttcaatatta	ttgaagcatt	tatcagggtt	attgtctcat	gagcggatac	atatttgaat	10500
gtatTTAGAA	aaataaaacaa	ataggggttc	cgcgcacatt	tcccccggaaa	gtgccacctg	10560
acgtctaaga	aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggc	10620
ccttcgtc						10629

<210> 29
<211> 676
<212> DNA
<213> Art

<220>
<223> Description of Artificial Sequence: 5' MAR in
Ar-Act-Af

```

<400> 29
ctccaccgcg gtggcgccg ctctagaact agtggatcct gcaagagttt tacaagtgaa 60
attaaggta aacattata aatatatacc aagtcaagt tgcttttaa aaattttgct 120
gacttcataa aatatattag gaagtcagct tggcatggag ctatataatc taaaaaagga 180
atcaaaaata tcgaaaaata tattagaagg gagaggacta attaatattt accagtcctaa 240
gggaaatata ttctttggag ctggcctatc gttttaaaat ttttttcat ggacgtcata 300
tataaatcta aacccttct aaatatatta gaaattaaag atatcgagat cctgcaagag 360
ttttacaagt gaaattaagt ttaaacattt ataaatatat accaagtcaa gtgtgttttt 420
taaaaatttt gctgacttca taaaatatat taggaagtca gcttggcatg gagctatata 480
atcttaaaaa ggaatcaaaa atatcgaaaa atatattaga agggagagga ctaattaata 540
tttaccagct ccagggaaat atattcttg gagctggct atcgttttaa aatttttttt 600
catggacgtc atatataaaat ctaaaccct tctaaatata ttagaaattt aagatatcga 660
gatccccccgg gctgca 676

```

```
<210> 30
<211> 15676
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence:pAct2-bin

```
<400> 30
gatctgcaga cgatcgcgcc gcgccattta aatgcggggc cgtttaaacgc cggccgcatt 60
ccccqqaagc ttgcatqcct qcagaattcg qcttattttq acqaaqttcgg atqtagtqatq 120
```

agccattatt taatgtacat actaatcgta
attgtataaa tatccataaa cacatcatga
tatgatacaa ttctaataga aaacgaatta
aacaaggcca ccacgacgac gactaacgtt
aaaaaaaaacg gagctgtcat gtaacacgcg
aaagcaaaag aactaatcca aggggctgag
cgagggaaaa ggctgcctga cagccaggtc
cgtgtctgtc gatttttaatt attttttga
cccgcctata taaaattcata tattttcctc
actttcatca gccgttttga atctccggcg
aagagagaaa gtaagagata atccaggaga
tcatcttctt ccgcttttc tttccaaggta
ctggatctcg atcttggtt ctcaatttcc
ctgtgaacct ccactaaatc ttttggttt
tagctcgatt atagctacca gaatttggct
tacctggaa atgatttga tatgtgaaat
ctgaacactg tcaatgttag attgaatctg
tatagattct tcgaaactt aggatttga
gattcaatca gggtttattt gactgtattt
catggccgt cctgttagaaa ccccaacccg
attcagtctg gatcgcgaaa actgtggat
agaaaagccgg gcaattgctg tgccaggcag
tcgttaattat gcgggcaacg tctgttatca
aggccagcgt atctgtctcg gtttcatgc
taatcaggaa gtgatggagc atcaggcgg
gtatgttattt gccggaaaaa gtgtacgtat
gcagactatc ccccccggaa tgggttattt
cttccatgtat ttcttaact atgcggaaat
gaacacctgg gtggacgata tcaccgttgt
gtctgttgac tggcagggtgg tggccaatgg
tcaacagggt gttgcaactg gacaaggcac
cctctggcaaa cccgggtgaag gttatcteta
agagtgtgat atctacccgc ttccgtcg
gttccttgatt aaccacaaac cgttctactt
cttacgtggc aaaggattcg ataacgtgt
gattggggcc aactcttacc gtacactcgca
ggcagatgaa catggcatcg tgggtattga
tttaggcatt ggtttcgaag cggcaacaa
caacggggaa actcagcaag cgcacttaca
aaaccaccca agcgtggta tgtggagtat
gcacggaaat atttgcac tggcggaaagc
cacctgcgtc aatgtatgt tctgogacgc
tgtgtctgtc ctgaaccgtt attacggatg
agagaagta ctggaaaaag aacttctggc
catcaccgaa tacggcgtgg atacgttagc
gagtgaagag tatcagtgtg catggctgga
cgccgtcgta ggtgaacagg tatgaattt
gcmcgtggc ggtaacaaga aaggatctt
ttttctgtctg caaaaacgct ggactggcat
caaacaatga atcaacaact ctcctggcgc
ttggagagct cgaatttccc cgatcgatc
aatcctgttg cccgtcttgc gatgattatc
gtaataatta acatgtatg catgacgtta
ccgcaattat acatttaata cgcgatagaa
ttatcgccg cgggtgtcatc tatgttacta
cttaattaag gccggcgtgc aggcatgaa
gcaccggta gtaatattgt acggctaaga
gcttctaat ttcaaactat tcgggcctaa
tatataaccgt tgtaatttga gctcggtgta
ttctgttgaa gtgcagccca tttcaccggta
aactgcccag gggaaagccat cttgagcgcg
aatagtgaat atgatgaaat attgtatctt
aagacactt cttcacggt ctgaattat
aattacgtt aattgtatga aatctaattt
gcctggattt actcggttta agttaaccac
gatcgagcag gtcacagtca tgaagccatc
atgattaattt agtttaaaaaa ttagttaaca
acgttatctt tacctgtggt cgaaatgatt
aaggccgaaa ataaagtgtt aagagataaa
tccgcttga attgtctcg tgcctccctc
acttgacaga gaagaacaag gaagaagact
ttcattctcc gtttgaatc ttccctcaatc
aataggaact ttctggatct actttatctt
ttgagatctg gaattcggtt aatttggat
actagaatcg atctaagttt accgatcagt
tgaccttgc gggagatcc atgttcatgt
gaaatctgaa ctgttgaagt tagatgaaat
aacactgttt aaggtagat gaagttgtg
gtgtcgtacg ttgaacagaa agctatttct
aactctttt gtgttgc agctgttagac
tgaaatcaaa aaactcgacg gcctgtggc
tgatcagcgt tggggaaaaa gcgcgttaca
ttttaacat cagttcgcgc atgcagatat
gcgcgaagtc tttataccga aaggttggc
ggtcaactcat tacggcaaaat tggggtcaa
ctatacgcctt tttgaagccg atgtcactgc
caccgtttt gtgaacaacg aactgaactg
cgacgaaaac ggcaagaaaaa agcagctta
ccatcgacg gtaatgtctt acaccacgg
gacgcgttc ggcgaagact gtaaccacgc
tgatgtcgc gttgaactgc gtgatgcgg
tagcggact ttgcaagtgg tgaatccgc
tgaactgtc gtcacagcga aaaggcagac
catccggtaa gtggcagtga agggcgaaca
tactggctt ggtcgtcatg aagatgcgg
gatggtgcac gaccacgcgtaatggactg
ttacccttac gctgaagaga tgctcactg
tgaactgtc gctgtcggt ttaacctctc
gccgaaagaa ctgtacagcg aagaggcagt
ggcgattaaa gagctgatag cgcgtgacaa
tgccaaacgaa cccgataccg gtcgcgaagt
aacgcgtaaa ctgcaccgcg cgcgtccgat
tcacaccgat accatcagcg atctcttgc
gtatgtccaa agcggcgatt tggaaacggc
ctggcaggag aaactgcata agccgattat
cgggctgcac tcaatgtaca ccgacatgt
tatgtatcac cgctctttt atcgctcag
cgccgatttt gcacctcgc aaggcatatt
caactcgacg cgcggaccg agtcggcggc
gaacttcggt gaaaaaccgc agcaggagg
accatcgatc gctacagcct cggtgggaa
aacatttgcg aataaagtttt cttaaagattt
atataatttc ttttgaattt cgttaagcat
tttatgagat gggttttt gattagagtc
aacaaaatat agcgcgcataa ctaggataaa
gatcgatcg gaaattgttta aacgcggccg
gctcgatct cacctacgat gggggcgtc
gcgaatttgg cctgttagacc tcaattgcga
cttttgggtt gatgtatgtc actggcagg
ataagtgcgt gttatgtttt gtttattt
caagtcggct agattgattt agccctgat
gaatggaaat ggttgcgtt gtacaacagag
ggatggaaat ggttgcgtt gtacaacagag

acgacagaac acccacggga ccgagcttcg caagcttgtt gtaactgaaa aaggaaaatt 3840
 attgtgccag gcagttgaaa gtcagcacct tttAACGAGT gctgaaatga cggctaaatg 3900
 ggaaacgtat ttaaaaaaaaa tcggtaaaag agaaggcaat caagagaact ttattacgaa 3960
 tatcaaaaaa ttcatgttc atttactgga agctgtaccc aacgatatag aaaaactaaa 4020
 ttttctgat taccaggaac agaaagaaaa agaagcagaa aaaagtattg taggaaaatg 4080
 tcctaagtgt ggcaacaata ttgtattaaa aaaatcgTT tatgggtgtt caaattatcc 4140
 tgaatgtaa tagtactttag ctgaacattt tagaaagaaaa aaactcacca aaacaaatgt 4200
 aaaagaatta ctagagggaa aagaaaccct ggtaaaagga atcaaaacga aagatagaaa 4260
 gtcctacaat gccgtgtaa aaatcgaga aaaggatAT attgatttt tatcttctc 4320
 aaaataaaca taaaagccct ttaaagaggg ctTTtatata ttaatcacaa atcacttac 4380
 acaaattcaca agtatttgt gattgttgc gataaaataa gaataagaag aaatagaaag 4440
 aagtgaatgtaa ttgtggaaa tttaggcgc caaaaagaaa aacgaaatga tacaccaatc 4500
 agtgcaaaaa aagatataat gggagataag acgggtcgT ttctgtctga cttgcaccat 4560
 atcataaaaa tcgaaacagc aaagaatggc ggaaacgtaa aagaagttat gaaataaga 4620
 cttagaagca aacttaagag tgtgttgc gtcagtgatc taaaatttt gtataatagg 4680
 aattgaagtt aaatttagatg ctaaaaattt gtaattaaga aggagtattt acatgaacaa 4740
 aaatataaaa tattctcaaa actttttaac gaatgaaaaa gtactcaacc aaataataaa 4800
 acaattgaat taaaagaaaa ccgataccgt ttacgaaattt ggacacaggta aaggcattt 4860
 aacgacgaaa ctggctaaaa taagtaaaca ggtAACGTCT attgaattt acgtcatct 4920
 attcaactta tcgtcagaaa aattttttactt gtaactcgt gtcaattttaa ttcaccaaga 4980
 tattctacag ttcaattcc ctaacaaaca gaggtataaa attgtggaa gtattccta 5040
 ccatttaagc acacaaattt ttaaaaaagt ggttttgcg agccatgcgt ctgacatcta 5100
 tctgattgtt gaaaaaaaaa aaaaaaaagcg taccttgcgat attcaccgaa cactagggtt 5160
 gctcttgcac actcaagtct cgattcagca attgcttaag ctggcagcgg aatgcttca 5220
 tcctaaacca aaagtaaaca gtgtcttaat aaaacttacc cgccatacc cagatgtcc 5280
 agataaatat tggaaagctat atacgtactt tgTTTCAAA tgggtcaatc gagaatatcg 5340
 tcaactgttt actaaaaatc agtttcatca agcaatgaaa cacccaaag taaacaattt 5400
 aagtaccgtt acttatgagc aagtattgtc tatttttaat agttatctat tatttaacgg 5460
 gaggaaataa gggctgagaa agcccagtaa gggaaacaact gtaggttcga gtcgcgagat 5520
 cccccggaaac caaaggaaagt aggtttaaacc cgctccgatc agggcggagcc acggcaggcc 5580
 gagaacattt gttccttgc ggcattcggg ttggcggatc aaacactaaa gctactggaa 5640
 cgagcagaag tcctccggcc gccagttgcg aggccgtaaa ggtgagcaga ggcacgggag 5700
 gttgccactt gccccgtcgc acgggtccga acggcatgga aaccggcccc gccaggcccc 5760
 ctgcgacgccc gacaggatct agcgtctgcg ttgggtcga cacaacacgc gccacgccc 5820
 cagttccgca aatagcccc aggaccgcca tcaatcgat cgggtcacct agcagagcgg 5880
 cagagatgaa caccgaccatc agcgcgtcga cagcgcctac cgtcgcccgcg acccgcccc 5940
 caggcggtag accgaaataa acaacaagct ccagaatagc gaaatattaa gtgcgcccag 6000
 gatgaagatg cgcattccacc agatccccgt tggaaatctgt cggacgatca tCACGAGCAA 6060
 taaacccgccc ggcaacgccc gcagcagcat accggcgacc cctcggcctc gctgttcggg 6120
 ctccacgaaa acggccggaca gatgcgcctt gtgagcgtcc ttggggccgt cctcctgttt 6180
 gaagaccgac agcccaatga tctcggccgc gatgtaggcg ccgaatgcca cggcatctcg 6240
 caaccgttca gcaacgcctt ccattggcTT ttctccctcg tgctcgaaa cggaccggaa 6300
 catctctgga gctttcttca gggccgacaa tcggatctcg cggaaatctt gcacgtcgcc 6360
 cgctccaaagc cgtcgaatct gagccttaat cacaattgtc aattttatc ctctgtttat 6420
 cggcagtctcg tagagcgcgc cgtgcggcccg agcgataactg agcgaagcaa gtgcgtcgag 6480
 cagtgcggcgc ttgttccgt aatgccagta aagcgcttgc tgctgaaccc ccagccggaa 6540
 ctgaccccac aaggcccttag cgTTTGCAT gcaccaggtc atcattgacc caggcgtgtt 6600
 ccaccaggcc gtcgcctcgc aactcttcgc aggcttgcgc gacctgtcg cgccacttct 6660
 tcacgcgggt ggaatccgat ccgcacatga ggcggaaagg ttccagctt agcgggtacg 6720
 gctcccggt cgagctgaaa tagtcgaaaca tccgtcggc cgtcggcgac agcttgcgg 6780
 acttctccca tatgaatttgc gtgttagtggt cggcagcaaa cagcacgacg atttctctgt 6840
 cgatcaggac ctggcaacagg gacgtttct tgccacggc caggacgcgg aagcggtgca 6900
 gcagcgcacac cgattccagg tgcccaacgc ggtcggacgt gaagcccatc gccgtcgcc 6960
 gttaggcgcga caggcattcc tcggccttcg tgtaataccg gccattgtac gaccagccca 7020
 ggtcctggca aagctcgtag aacgtgaagg tgatcggtc gccgataggg gtgcgttcg 7080
 cgtactccaa cacctgtgc cacaccagtt cgtcatcg cggccgcgcg tcgacgcccc 7140
 ttaggtgtat ttcatgttc ttgttgcgt ggaaatgac ctgttttgc agcgcctcgc 7200
 gccccgatTTT ctgttgcgc gtggtaaca gggcagagcg ggcgtgtcg ttggcatcg 7260
 ctcgcacatgt gtccggccac ggcgcaatTTT cgaacaaagg aagctgcatt tccttgatct 7320
 gctgcttcgt gtgtttcagc aacgcggcct gcttgcctc gctgacctgt ttggccaggt 7380
 ctcgcggcgc ggttttgcg ttcttgcg tcatagttcc tcgcgtgtcg atggcatcg 7440

acttcgccaa acctggcgcc tcctgttcga gacgacgcga acgctccacg gcggccatgt 7500
gcccggcag ggcaaaaaaa gccagtttgcg cgtgtcgca ctcatcttg gccgttagct 7560
gctggaccat cgagccgacg gacttggaaagg tttcgccccg cgcacgcatt acgggtcgcc 7620
ttgcgtatgtt ttccggatcc tcggcgaaaa acccccggtc gatcgttct tgccctgtat 7680
ccttcggtc aaacgtccga ttccatttacc ctccttgcgg gattggcccg actcacgcgg 7740
gggcaatgtg cccttattcc tgatttggacc cgcctgggtc cttgggtgtcc agataatcca 7800
ccttatcgcc aatgaagtgc gtcccgtaga cgcgtctggcc gtccttctcg tacttggat 7860
tccgaatctt gccctgcacg aataccagcg acccccttgcg caaataacttg ccgtggccct 7920
cgccctgaga gccaaaacac ttgatgcgg aagaagtcgg ggcgccttcgc ttgtcgccgg 7980
catcggtcgcc caactcttca ttaaccgcta tatcgaaaaat tgcttgcggc ttgttagaat 8040
tgccatgacg tacccctgggt tcacgggtaa gattaccgat aaactggAAC tgattatggc 8100
nnctgaaat tcctcggtc ttgccttgct cgtcggtgat gtacttcacc agctccgcg 8160
agtcgtctt ctgtatggag cgcatggga cgtgcttgc aatcacgcgc acccccccggc 8220
cgtttagcg gctaaaaaaag tcatggctct gccctcgggc ggaccacgccc catcatgacc 8280
ttgccaagct cgtcctgtct ctcttcgatc ttgcggcagca gggcgaggat cgtggcatca 8340
ccgaaccgcg cctgtcgccgg gtcgtcggtg agccagagtt tcagcaggcc gcccaggcgg 8400
cccaggcgcg cattgtatgcg ggcacgtcg cggacgtgct catagtcac ggcgcggcc 8460
attttgtac cctggccgac ggcacggcagg taggcccggaca ggctcatgcc ggcgcggcc 8520
gcctttcct caatcgctct tcggtcgctt ggaaggcagt acacccctgt aggtggctg 8580
cccttcctgg ttggcttggg ttcatcagcc atccgttgc cctcatctgt tacggccggc 8640
gtagccggcc agcctcgacg agcaggattc ccgttgagca cgcgcagggtg cgaataagg 8700
acagtgaaga aggaacacccc gtcgcgggt gggcctactt cactatctt gcccgcgt 8760
cgccgttggg tacaccaagg aaagtctaca cgaacccctt ggcaaaatcc tgtatatcgt 8820
gcaaaaaagg atggatatac cggaaaaatc gctataatga cccgaagca gggttatgca 8880
gccccggaaa tecgtcgacc ctttccgacg ctcaccggc tgggtgcctt cggcgttggg 8940
ctggcgccg tctatggccc tgcaaacgcg ccagaaacgc cgtcgaaagcc gtgtcgaga 9000
caccgcggcc ggccggccggc gttgtggata ccacgcgaa aacttggccc tcactgac 9060
atgaggggcg gacgttgaca cttgaggggc cgactcacc ggcgcggcgt tgacagatga 9120
ggggcaggct cgatttcggc cggcgacgtg gagctggca gcctcgaaa taagtggccct 9180
cgccctgatt tacgcgaggat tcccacagat gatgtggaca agcctgggggtaaagtgccct 9240
gccccgttga cacttgaggg ggcgactac tgacagatga gggcgcgat ccttgacact 9300
tgagggcggcag agtgcgtaca gatgaggggc gcacatttgc acatttgagg ggtgttccac 9360
aggcagaaaa tccagcattt gcaagggttt ccgcccgtt ttcggccacc gctaacctgt 9420
cttttaacct gcttttaaac caatatttt aaaccttgc ttaaccagg gctgcgcct 9480
ggcgcgtgac cgccacgcgc gaaggggggt gccccccctt ctogaaccct cccggccgc 9540
taacgcgggc ctcccatccc cccaggggct gcgccccctcg gccgcgaacq gcctcacccc 9600
aaaaatggca ggccaagctt gcttggcgat tccgcgtgaa cgtcggtcg attgtaacctg 9660
cgttcaataa cttgcgatc gtgttgcgat cctgcggcgat gctcggtcg atctcacg 9720
tcgactgttt ctctcgcaac gccatccgac ggtatgtt taaaagtccc atgtggatca 9780
ctccgttgcc ccgtcgctca ccgtgttggg gggagggtgc acatggctca gttctcaatg 9840
gaaattatct gcctaaccgg ctcagttctg cgtagaaacc aacatgcaag ctccacccgg 9900
tgcaaagccg cagcgccggc aggatatatt caattgtaaa tggcttcatg tccggaaat 9960
ctacatggat cagcaatgag tatgtatggc aatatggaga aaaagaaaga gtaattacca 10020
atttttttc aattcaaaaaa tgcgtatgtc cgcagcgat ttataaaatg aaagtacatt 10080
ttgataaaac gacaaattac gatccgtcg atttataaggc gaaagcaata aacaaattat 10140
tctaattcgg aaatctttat ttgcgtatgt ctacatttac gtcggatcc gggagatccg 10200
tcgacctgca gtgagggtctg cctcgatgg aagggtgtgc tgactcatac caggcctgaa 10260
tcgccccatc atccagccag aaagtgggg agccacgtt gatgagagct ttgtttagg 10320
tggaccagtt ggtgatttt aactttgtt ttgcacggc acgtctcgat ttgtcgccgg 10380
gatgcgtat ctgatccctt aactcagcaa aagttcgatt tattcaacaa agccacgtt 10440
tgtctcaaaa tctctgtatgat tacattgcac aagataaaaa tatatcatca tgaacaataa 10500
aactgtctgc ttacataaaac agtaatacaa ggggtgttat gacccatatt caacggaaa 10560
cgcttgcgc gaggccgcga ttaaatttca acatggatgc tgatttataat gggtataaaat 10620
gggctcgca taatgtcgaa caatcaggtg cgacaaatcta tcgattgtat gggagcccg 10680
atgcgcaga gttgttctg aaacatggca aaggtagcgt tgccaatgtat gttacagat 10740
agatggtcag actaaactgg ctgacggat ttatgccttcc gccgaccatc aagcatttta 10800
tccgtactcc tgcgtatgtca tgggtactca ccactcgat ccctggggaaa acagcattcc 10860
aggtattaga agaataatcc gattcaggtg aaaatattgt tgatgcgtcg gcaagtgttcc 10920
tgcgcgggtt gcattcgatt cctgttgcgat ttgtcctttaa acagcgtat gcgcttatttc 10980
gtctcgctca ggcgcaatca cgaatgaata acgggttgggt tgatgcgtat gatttgtat 11040
acgagcgtaa tggctggccct gttgaacaag tctggaaaga aatgcataaa cttttgccat 11100

tctcaccgga ttcagtcgtc actcatggtg atttctcaact tgataacacctt atttttgacg 11160
 agggaaaatt aatagggttgtt attgtatgtt gacgagtcgg aatcgccagac cgataccagg 11220
 atcttgcac cctatggaaac tgcctcggtg agtgttctcc ttcatcatacg aaacggcttt 11280
 ttcaaaaata tggtattgtt aatcctgata tgaataaattt gcagtttcat ttgtatgctcg 11340
 atgagttttt ctaatcagaa ttggtaattt gggtgttaaca ctggcagagc attacgctga 11400
 cttgacggga cggcggctt gttgaataaa tcgaactttt gctgagttga aggatcagat 11460
 cacgcacatctt cccgacaacg cagaccgttc cgtggcaaag caaaagttca aaatcaccaa 11520
 ctggccacc tacaacaaaag ctctcatcaa ccgtggctcc ctcactttct ggctgatgat 11580
 tggggcgatt caggcctgtt atgagtcaac aacaccttct tcacgaggca gacctcagcg 11640
 cctgcaggc gacggatctg ggggatctg cagatccgca agggatcga gcccacata 11700
 tgccccgtt tcggtgcgac taacatgagt tcttggacaa atttattttg acctgatgag 11760
 atgatccaac ccgaggatattt agcaaagctc gttcgtgcag caatgaaacg gccaaaccgt 11820
 gctttgtcc ccaagaatga ggtqctatgc atgaaggaaat ctacccgtt atgtccaaca 11880
 gtctcagggt taatgtctat gtatctaaa taatgttgc ggtatttgtt aatctcatat 11940
 agatttcac tgcgtgcacgca aaaaatatta aataaatattt attattatct acgtvttgat 12000
 tgagatataca tcaatattat aataaaaata tccattaaac acgattttgat acaaatagaca 12060
 gtcaataatc tgatttgaat atttattat tctaacaat tacataaaaga tcgaatagaa 12120
 aatactgcac tgcaaatgaa aattaacaca tactaataaa tgctcaaat atcttgcca 12180
 agatcaagcg gagtgaggc ctcataatccg gtctcagttt caagcacggg atccccgaag 12240
 cgcgcctcac caatgcctc gacatagatg ccgggctcga cgttggaggac attgcctacc 12300
 ttgagcatgg tctcagcgcc ggctttaagc tcaatcccat cccaatctga atatcttac 12360
 ccgcggccag tccgggttaa gaacgggtct gtccatccac ctctgttggg gtggcgaag 12420
 aactccagca tgagatcccc gcgcgtggagg atcatccagc cgcgtcccg gaaaacgatt 12480
 ccgaagccca acctttcata gaaggcggcg gtggaaatcga aatctcgtga tggcagggtt 12540
 ggcgtcgctt ggtcggtcat ttgcgaccccc agagtcccgc tcagaagaac tcgtcaagaa 12600
 ggcgatagaa ggcgatgcgc tgcaatcg gacggcgat accgtaaacg acgaggaagc 12660
 ggtcagccca ttgcggcaca agctttcag caatatcagc ggtagccaaac gctatgtcct 12720
 gatagcggcgc cggccacaccc agccggccac agtcgatgaa tccagaaaaag cggccatttt 12780
 ccaccatgtt attcggcaag caggcatcgc catgggtcac gacgagatcc tcggcgtcgg 12840
 gcatgcgcgc cttgagccctg gcaaacagttt cggctggcgc gagccccctga tgctttcg 12900
 ccagatcatac ctgatcgaca agaccggctt ccatccagat acgtgctcgc tcgatgcgt 12960
 gtttcgtt gttgtcaat gggcaggtag ccggatcaag cgtatgcagc cggccatttg 13020
 catcagccat gatggataact ttctcgccag gagcaagggtt agatgacagg agatcctgcc 13080
 ccggcacttc gcccaatagc agccagtttcc ttcccgttcc agtgacaacg tcgagcacag 13140
 ctgcgcagg aacggccgtc gtggccagcc acgatagccg cgtgcctcg tcctgcagtt 13200
 cattcaggcgc accggacagg tcggcttgc aaaaaaaac cggccggcccc tgcgtgaca 13260
 gccggAACAC ggcggcatca gagcagccga ttgtctgtt tgcccaatc tagccgaata 13320
 gcctctccac ccaagcgccgg ggagaacctg cgtcaatcc atcttgcata atcatgctaa 13380
 aggatctcga tccccgggtt gattttctca gtctccagag atgtgtttaa ataggcagta 13440
 gcctttgtat atcagccaca agtgtgtggg aatcttatct tcggatttca attaggaatt 13500
 aaccttattt aattctctt aagggaaatgc cgcaaaatgtt ttgtctgtt ccattatgtt 13560
 ttcaacatca aatgaatagt ggctaagcc tgcttgcattt ctgatgtttc ttccaagtt 13620
 cgaatctcct ttgtatattt gattaacgaa actcttggaaa ttagtattat ctgtcttaat 13680
 cagaaaatga acaggagttt gataaataact aaatttcttt atagtattttt ttacogccaa 13740
 tgtctcttttgcattgtgtt ggttaattttt ttctgcagttt ttaaagctca gatcttgcag 13800
 gtccccagat tagccttttcaatttccgaa agaatgttcc cccacagatgg gtttagagagg 13860
 cttacgcagc aggtctcatc aagacgttcc acccgagccaa taatctccag gagatcaaat 13920
 acctttccaa gaagggtttttt gatgcgtca aaagatttccg gactaactgc atcaagaaca 13980
 cagagaaaga tatatttctc aagatcgaa gtactattcc agtatggacg attcaaggct 14040
 tgcttcacca accaaggccaa gtaatagaga ttggagttcc taaaaaggta gttccactg 14100
 aatcaaaggc catggagttca aagatttccaa tagaggaccc aacagaactc gccgtaaaga 14160
 ctggccaaac gttcatacag agtcttttac gactcaatga caagaagaaaa atcttcgtca 14220
 acatgggttca gacgcacacg cttgtctact caaaaatata caaagataca gtctcagaag 14280
 accaaaggccaa aatttggactt tttcaacaaa gggtaatattt cggaaacccctc ctcggattcc 14340
 attgcccagc tatctgttccat tttattttttt gatagttttt gatagttttt ggctccatca 14400
 aatgcctatca ttgcataaaa ggaaaggccaa tcgttgcataa tgccctctgc gacagtggc 14460
 ccaaagatgg acccccaccc acggaggagca tcgtggaaaa agaagacgtt ccaaccacgt 14520
 cttcaaaagca agtggattttt gttgtatatttcc actatcccttcc gcaagacccctt 14580
 aatccatccatca ttgcataaaa ggaaaggccaa tcgttgcataa tgccctctgc gacagtggc 14640
 gggggacttcc agaggatccg gatccgtca ccatgagccaa aggagaagaa cttttctactg 14700
 gagttgttttttttcc aatttgcattttt gatagttttt gatagttttt ggctccatca 14760

gtggagaggg	tgaagggtat	gctacatacg	gaaagcttac	ccttaaattt	atttgcaact	14820
ctggaaaact	acctgttcca	tggccaaacac	tttgtcaactac	tttctcttat	ggtgttcaat	14880
gctttcccg	ttatccggat	catatgaagc	ggcacgactt	cttcaagagc	gccatgcctg	14940
agggataacgt	gcaggagagg	accatcttt	tcaaggacga	cgggaactac	aagacgcgtg	15000
ctgaagtcaa	gtttgaggga	gacaccctcg	tcaacaggat	cgagcttaag	ggaatcgatt	15060
tcaaggagga	cgaaaacatc	ctcgccccaca	agtttgaata	caactacaac	tcccacaacg	15120
tatacatcac	ggcagacaaa	caaaaagaatg	gaatcaaagc	taacttcaaa	attcgccaca	15180
acattgaaga	tggatccgtt	caacttagcag	accattatca	acaaaatact	ccaattggcq	15240
atggccctgt	ccttttacca	gacaaccatt	acctgtcgac	acaatctgcc	ctttcgaaag	15300
atcccaacga	aaagcgtgac	cacatggtcc	ttcttgagtt	tgtaactgtct	gctgggatta	15360
cacatggcat	ggatgagcta	tacaaataaa	gatcctcgaa	tttccccgat	cgttcaaaca	15420
tttggcaata	aagtttctta	agattgaatc	ctgttgcgg	tcttgcgtatg	attatcatat	15480
aatttcgtt	gaattacgtt	aagcatgtaa	taattaacat	gtaatgcattg	acgttattta	15540
tgagatgggt	ttttatgatt	agagtcccg	aattatacat	ttaatacgcg	atagaaaaaca	15600
aaatatagcg	cgcaaactag	gataaattat	cgcgcgcgg	gtcatctatg	ttactagatc	15660
qatcgggaat	tagatc					15676

<210> 31

<211> 17111

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:pArActAf-bin

<400> 31

gatctgcaag	cgatcgccggc	gcgccattta	aatgcccggg	cgtttaaacg	cggccgcctct	60
agaactagt	gatcctgcaa	gagtttaca	agtgaattta	agtttaaca	tttataaaata	120
tataccaagt	caagtgtgt	ttttaaaaat	tttgctgact	tcataaaaata	tatttaggaag	180
tcagcttggc	atggagctat	ataatcttaa	aaaggaatca	aaaatatcg	aaaatatattt	240
agaagggaga	ggactaatta	atatttacca	gctccagggg	aatatattct	ttggagctgg	300
cctatcgtt	taaaattttt	tttcatggac	gtcatatata	aatctaaacc	ccttcttaat	360
atattagaaa	ttaaaagatat	cgagatcctg	caagagttt	acaagtgaaa	ttaagttaa	420
acatttataa	atatataacca	agtcaagtgt	gctttttaaa	aattttgctg	acttcataaa	480
atatatattgg	aagtcaagctt	ggcatggagc	tatataatct	taaaaaggaa	tcaaaaatat	540
cgaaaaatat	attagaaggg	agaggactaa	ttaatattta	ccagctccag	ggaaaatatat	600
tctttggagc	tggcctatcg	ttttaaaatt	tttttcatg	gacgtcatat	ataaatctaa	660
accccttcta	aatatatttag	aaattaaaga	tatcgagatc	ccccgggctg	cagaatcg	720
cttggtttga	cgagttcgga	tgttagtagta	gccatttattt	aatgtacata	ctaatcgta	780
atagtgaata	tgtgaaaca	ttgtatctta	ttgtataaaat	atccataaaac	acatcatgaa	840
agacacttcc	tttcacggtc	tgaatttaatt	atgatacaat	tctaataagaa	aacgaattaa	900
attacgttga	attgtatgaa	atctaattga	acaagccaaac	cacgacgacg	actaacgttg	960
cctggattga	ctcggtttaa	gttaaccact	aaaaaaaacgg	agctgtcatg	taacacgcgg	1020
atcgagcagg	tcacagtcat	gaagccatca	aagccaaaaga	actaatccaa	ggggctgaga	1080
tgattaatta	gtttaaaaat	tagttaacac	gagggaaaag	gctgcctgac	agccaggtca	1140
cgtttatctt	acctgtggtc	gaaatgattc	gtgtctgtcg	attttaatta	tttttttgaa	1200
aggccgaaaa	taaagttgt	agagataaaac	ccgcctatata	aaattcatat	attttcctct	1260
ccgcttggaa	ttgtctcggt	gtcctccctca	ctttcatcag	ccgttttggaa	tctccggcga	1320
cttgcacagag	aagaacaagg	aagaagacta	agagagaaaag	taagagataa	tccaggagat	1380
tcattctccg	tttgaatct	tcctcaatct	catcttcttc	cgctctttct	ttccaaggta	1440
ataggaactt	tctggatcta	ctttatttgc	tggatctcg	tcttttttc	tcaatttcct	1500
tgagatctgg	aattcgttt	atttgggatc	tgtgaacctc	cactaaatct	tttggtttta	1560
ctagaatcga	tctaagttga	ccgatcagtt	agctcgatta	tagtaccag	aatttgctt	1620
gaccttgatg	gagagatcca	tgttcatgtt	acctggggaa	tgatttgtat	atgtgaaatg	1680
aaatctgaac	tgttgaagtt	agattgaatc	tgaacactgt	caatgttaga	ttgaatctga	1740
acactgttta	aggttagatg	aagtttgtt	atagattctt	cgaaaacttta	ggattttgtag	1800
tgtcgtaacgt	tgaacagaaaa	gctatttctg	attcaatcag	ggtttatttg	actgtattga	1860
actctttttg	tgtgtttgca	gctgttagacc	atggtcgtc	ctgtagaaac	cccaaccgt	1920
gaaatcaaaa	aactcgacgg	cctgtggca	ttcagtcctgg	atcgcgaaaa	ctgtggaatt	1980
gatcagcggt	ggtggggaaag	cgcgttacaa	gaaagccggg	caattgtgt	gccaggcagt	2040
tttaacgatc	agttcgccga	tgcagatatt	cgtattatg	cggcaacgt	ctggtatcag	2100

cgcaaggct ttataccgaa aggttggcga ggccagcgta tcgtgctgcg tttcgatgcg 2160
 gtcactcatt acggcaaagt gtgggtcaat aatcaggaag tcatggagca tcagggcgcc 2220
 tatacgccat ttgaagccgaa tgtcacgcg tatgttattt ccggggaaaag tgcgtatc 2280
 accgttgtg tgaacaacga actgaactgg cagactatcc cgccggaaat ggtgattacc 2340
 gacgaaaacg gcaagaaaaa gcagtcttac ttccatgatt tctttaacta tgccgaaatc 2400
 catcgacgca taatgctcta caccacgcg aacacctggg tgacgatac caccgtggg 2460
 acgcacgtcg cgcaagactg taaccacgcg tctgttact ggcaggtgg ggcacatgg 2520
 gatgtcagcg ttgaactgcg tgatgcggat caacagggtgg ttgcaactgg acaaggcact 2580
 agcgggactt tgcaagtgtt gaatccgcac ctctggcaac cggtgaaagg ttatcttat 2640
 gaactgtgcg tcacagccaa aagcagaca gagtgtgata tctaccgcg tgcgtcgcc 2700
 atccggtcag tggcagtggaa gggcgaacag ttccctgatta accacaaacc gttctacttt 2760
 actggctttg gtcgtcatga agatgcggac ttacgtggca aaggattcga taacgtctg 2820
 atggtgacg accacgcatt aatgactgg attggggcca actcctaccg tacctcgcat 2880
 tacccttacg ctgaagagat gctcgactgg gcagatgaac atggcatcgt ggtgattgat 2940
 gaaactgctg ctgtcggtt taacctctct ttaggcattt gtttgcggc gggcaacaag 3000
 ccgaaaagaac tgcgtacgcg agaggcagtc aacggggaaa ctcagcaacg gcacttacag 3060
 gcgattaaag agctgtatgc gcgtgacaaa aaccacccaa gcgtgggtat gtggagtatt 3120
 gccaacgaaac cggataccgc tccgcacgtg cacgggaaata tttcgccact ggcggaaagca 3180
 acgcgtaaac tcgacccgac gcgtccgatc acctgcgtca atgtaatgtt ctgcgacgct 3240
 cacacccata ccatcagcga tctcttgcgat gtgcgtgccc tgaaccgtta ttacgatgg 3300
 tatgtccaaa gcggcgattt gaaacggca gagaaggatc tggaaaaaga acttctggcc 3360
 tggcaggaga aactgcatac gccgattatc atcaccggaaat acggcgtggaa tacgttagcc 3420
 gggctgcact caatgtacac cgacatgtgg agtgaagagt atcagtgtgc atggctggat 3480
 atgtatcacc gcgtctttga tcgcgtcagc gccgtcgctg gtgaacaggat atggatttc 3540
 gccgattttg cgacacgcgcg aggcatattt cgcgtggcg gtaacaagaa agggatcttc 3600
 actcgcgacc gcaaaccgaa gtcggcgct tttctgcgtc aaaaacgcgt gactgcgt 3660
 aacttcggtg aaaaaccgca gcagggaggc aaacaatgaa tcaacaactc tcctggcgca 3720
 ccatcgctgg ctacagccctc ggtggggaaat tggagagctc gaatttcccc gatcggtcaa 3780
 acatttggca ataaagtttca ttaagattga atcctgttgc cggttgcgt atgattatca 3840
 tataatttct gttgaattac gttaaagcatg tataattaa catgtaatgc atgacgttat 3900
 ttatgagatg gttttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 3960
 acaaaaataa ggcgcgaaac taggataatatt tatcgccgatc ggtgtcatct atgttactag 4020
 atcgatcggg aattgtttgg gggatctcgat tatctttat ttcataatata tttagaaagg 4080
 gtttagattt atatatgcg tccatgaaaaaa aaaaattttaa aacgataggc cagctccaaa 4140
 gaatataattt ccctggagat ggttaatatt aattagtcct ctccttcta atatattttt 4200
 cgatattttt gattcctttt taagattata tagctccatg ccaagctgac ttcctaataat 4260
 attttatgaa gtcagaaaaa tttttaaaaaa gcacacttgc ctggatataat atttataat 4320
 gtttaaactt aatttcaccc taaaactct tgccaggatct cgatatctt aatttctaat 4380
 atatttagaa ggggtttaga ttatataatg acgtccatga aaaaaaattt taaaacgata 4440
 ggcctggctcc aaaaatataa tttccctggaa gctggtaat attaattatg ctcctccctt 4500
 ctaatataattt tttcgatattttt tttgattccct ttttaagattt atatagctcc atgcctaaagct 4560
 gacttcctaa tatattttat gaagtcagca aattttttaa aacgcacact tgacttggta 4620
 tatatttata aatgtttaaa cttaatttca cttgtaaaac tttgcaggaa tccactatgtt 4680
 ctagagcggc cgcttaatattt aggccggctt gcaggcatgc aagctcggt ctcacccatcg 4740
 atggggggca tcgcaccggat gagaatattt gtacggctaa gacgtggaaat ggcctgtaga 4800
 cctcaattgc gagcttttca atttcaaactt attcggccctt aactttgtt gtgatgtatgc 4860
 tgacttggcag gatataacc gttgtatattt gagctcgatc gataaagtgc ctgtgtatgt 4920
 ttgtttgattt gttctgttgc gagtgcagcc catttcaccgc gacaaggatcg ctagattgt 4980
 tttagccctga tgaactgcgcg agggaaagcc atcttgagcg cgaatggga atggatttcg 5040
 ttgtacaacg agacgcacaga acacccacgg gaccgagctt cgcacgttgc ttgtactga 5100
 aaaagaaaaa ttattgtgcc aggcatgtt aagtgcac cttttaaacga gtgctgaaat 5160
 gacggctaaa tgggaaacgt attttaaaaaa aatcggtaaa agagaaggca atcaagagaa 5220
 ctttattacg aatataaaaaa aattcattgt tcatttactg gaagctgtac ctaacgatata 5280
 agaaaaacta aattttctg attaccaggaa acagaaagaa aagaaggacg aaaaagat 5340
 tgttagaaaaa tgcgtcaatgtt gtcgtggaa tattgttattt aaaaatcgat tttatggat 5400
 ttcaattat cctgaatgtt agtttactttt agctgaacat tttagaaaga aaaaactcac 5460
 caaaaacaaat gtaaaaagaat tactagaggg aaaaagaaacc ctggtaaaag gaatcaaaac 5520
 gaaagataga aagtccatca atgcgttgcg aaaaatcgaa gaaaaggat atattgattt 5580
 tatatcttcc tcaaaaataaa cataaaagcc cttaaagag ggctttata tattatcaca 5640
 aaatcactt tcaaaaaatca caagtgttgc gttgttgcg atgataaaaat aagaataaga 5700
 agaaatagaa agaagtgttgc gttgttgcg aatttaggcg cacaaaaaa aaaaacgaaat 5760

gatacaccaa tcagtgc当地 3aaagatata atgggagata agacggttcg tgttcgtgct 5820
 gacttgc当地 atatcataaa aatcgaaaaca gcaaagaatg gccggaaacgt aaaagaagtt 5880
 atggaaataa gacttagaag caaacttaag agtgtgtga tagtgc当地 tcttaaaatt 5940
 ttgtataata gaaattgaag ttaaattaga tgctaaaaat ttgttaattaa gaaggagtga 6000
 ttacatgaac aaaaatataa aatattctca aaactttta acgagtgaaa aagtactcaa 6060
 ccaaataata aaacaatgaa attaaaaga aaccgataacc gtttacgaaa ttggaacagg 6120
 taaagggcat ttaacgacga aactggctaa aataagtaaa caggtaacgt ctattgaatt 6180
 agacagtcat ctattcaact tattcgtcaga aaaattaaaa ctgaataactc gtgtcactt 6240
 aattcaccaa gatattctac agttcaatt ccctaacaaa cagaggata aaattgttgg 6300
 gagtattcct taccatgaa gcacacaaaat tattaaaaaa gtggtttttgg aaagccatgc 6360
 gtctgacatc tatctgatttgaagg attctacaag cgtagcttgg atattcaccg 6420
 aacactaggg ttgtcttgc acactcaagt ctcgatttag caattgctta agtgc当地 6480
 ggaatgctt catcctaaac caaaagtaaa cagtgctta ataaaactta cccgc当地 6540
 cacagatgaa ccagataaat attggaagct atatacgtac tttgttcaa aatgggtcaa 6600
 tcgagaatat cgtcaactgt ttactaaaaa agtaaacaat ttaagtaccg ttacttatgaa attatttaac gggagggaaat aattctatgaa ttactaaagg gaatgttagat aaattattag agaggccct acgc当地tacg gggaaatttgg gctc当地ggc tgagaaagcc cagtaaggaa cggaaacaaa ggaagtaggt taaacccgct acattggctt ctgttaggcat cgggattggc agaagtccct cggccgccaag ttgccaggcc cacttgc当地 ggtagaccgat ccaccagatt cccggttggaa agatgc当地 catcctccatg gatctagcgc tgcgttggt ccccaatag ccccccaggac cgccatcaat atgaacacga ccatcagcgg ctgc当地cagcg ggtagaccgat aataaacaac aagctccaga agatgc当地 catcctccatg gatctagcgc tgcgttggt gttc当地cgaa cgc当地tccatg ggcttttctt ctggagctt cttcaggccc gacaatcggaa caagccgtcg aatctgagcc ttaatcacaat gttc当地tagag cgc当地cggtc gccc当地gagca cccgcttggt cctgaaatgc cagtaaaggc cccacaaggc cctagcgtt gcaatgc当地 agggccgtcg ctc当地caactc ttgc当地aggct cggg当地ggaaat ccgatccgca catgaggccg cggtgc当地 gtaaaatagc gaacatccgt tcccatatga atttcgtgtt gttgtc当地 gcaaaacagca aggtcatcat tgacccaggc gtttccacc 8040
 aggacttggc aacgggacgt tttcttgc当地 cggg当地ggaaat ccgatccgca catgaggccg cggtgc当地 gtaaaatagc gaacatccgt tcccatatga atttcgtgtt gttgtc当地 gcaaaacagca aggtcatcat tgacccaggc gtttccacc 8040
 gacaccgatt ccagggccccc aacgc当地ggcg gacgtgaaagc ccatcgc当地 cccctgttggt cctcgtgtt gacgtggaaa atgaccttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 cgc当地acaggc atttcctccggc cttcgtgtt gacgtggaaa atgaccttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 tggcaagact cgttagaactg gaaggtgatc ggctc当地ggca taggggtgctc ttccgtgtt gacgtggaaa atgaccttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 tccaaacacct gctgccc当地 acacttgc当地 cttcgtgtt gacgtggaaa atgaccttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 gtgatctca cgtcccttggt gacgtggaaa atgaccttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 attttcttgc当地 tgc当地gtgtt gacaggggca gagcgggccc tgc当地gttgg catcgtctcg 8700
 atcgtgtccg gccacggccgc aatatcgaac aaggaaagct gcatcgttgc当地 ttc当地acccat cagttcgtca tcgtc当地ggcc gacgtcgac gccggtgttag 8400
 ttc当地gtgtt tcagcaacgc ggc当地gttgc当地 gcttc当地gttgc当地 cctgttttgc当地 caggttccctcg 8820
 cccggccgtt ttc当地gttgc当地 ggtc当地gttgc当地 gttc当地gttgc当地 tgc当地gttgc当地 catcgtactt 8880
 gcca当地acctg cc当地gttgc当地 ttc当地gttgc当地 cggc当地gttgc当地 cggc当地gttgc当地 ccatcggccgc cgtggccgc 8940
 ggc当地aggccag ggggaggccag ttgc当地gttgc当地 tgc当地gttgc当地 tcttggccgt agcttgc当地 9000
 accatcgagc cgc当地ggactg gaagggttgc当地 cggggccgc当地 gcatgacggt gccggtgtcg 9060
 atggtttccgg catcctccggc ggaaaacccc gctgtc当地gttgc当地 gttc当地gttgc当地 tgc当地gttgc当地 gcatgacggt gccggtgtcg 9060
 cggtcaaacgc tccgattcat tcacccttgc当地 tgc当地gttgc当地 cggggtatttgc当地 ccccgactca cggccggggca 9180
 atgtgccc当地tatttccgtt gacccc当地gttgc当地 ggtc当地gttgc当地 tgc当地gttgc当地 tgc当地gttgc当地 catcgtactt 9240
 tcggcaatga agtc当地gttgc当地 gtagaccgtc tggccgttgc当地 tctcgtactt ggtattccga 9300
 atcttgc当地ccctt gcaacatc cagc当地ggcccc tgc当地gttgc当地 acttgc当地gttgc当地 ggc当地gttgc当地 ggc当地gttgc当地 9360
 tgagagccaa aacacttgc当地 gcgaaagaag tgc当地gttgc当地 cctgttgc当地 ggc当地gttgc当地 ggc当地gttgc当地 9420

ttgcgccact cttcattaac cgctatatacg aaaattgctt gcccgttgtt agaattgcc 9480
 tgacgtacct cgggtcacg gtaagatta ccgataaaact ggaactgatt atggcnnc 9540
 gaaattccct cggcttgcgc ttgcgtcg gtatgtact tcaccagctc cgcgaagtgc 9600
 ctcttcttga tggagcgcatt ggggacgtgc ttggcaatca cgcgcacccc cccggcg 9660
 tagcgcctaa aaaagtcatg gctcgccct cgggcggacc acgcccatac tgaccttgcc 9720
 aagctcgcc tgccttcattt cgatsttcgc cagcaggcg aggatcggtt catcaccgaa 9780
 ccgcgccgtg cgcgggtcg cggtgagcca gagttttagc aggcggccca ggccggcc 9840
 gtcgcattt atgcgggcca gctcgcggac gtgcctatag tccacgacgc ccgtat 9900
 gtagccctgg cgcacggcca gcaggtaggc cgacaggctc atgcggccg cccggcc 9960
 ttccctaattt gctttcggtt cgtcgaaag gcaatcacc ttgataggtt ggctggcc 10020
 cctgggtggc ttggtttcat cagccatccg cttggccctca tctgttacgc cggcgtagc 10080
 cggccagcc cgcagagcag gattcccggt gaggccccc agtgcgaat aaggacagt 10140
 gaagaaggaa cacccgctcg cggggggcc tacttcaccc atccctgccc gctgacgccc 10200
 ttggatacac caaggaaagt ctacacgaa cctttggcaa aatcctgtat atcggtcgaa 10260
 aaaggatgga tataccgaaa aaatcgctat aatgaccccg aagcagggtt atgcagcg 10320
 aaagatccgt cgacccttc cgacgctcac cgggctgggtt gcctcgcc cttggctggc 10380
 gggcgcttat ggccctgcaaa acgcgc 10440
 cggccggccg cggcggttgtt ggataccacg cggaaaactt gcccctcact gacagatgag 10500
 gggcgacgt tgacacttga ggggscgact caccggcgc ggcgttgaca gatgaggggc 10560
 aggctcgatt tcggccggcg acgtggagct ggccaggctc gcaaatcgcc gaaaacgc 10620
 gatttacgc gagtttccca cagaatgtt gacaaggctt gggataagt gcccctggc 10680
 attgacactt gaggggcg actactgaca gatgaggggc gcgatcctt acacttgagg 10740
 ggcagagtga tgacagatga ggggscgacc tattgacatt tgaggggctg tccacaggca 10800
 gaaaatccag catttgc 10860
 aacctgc 10920
 gtgacccgc acgccc 10980
 cgggcctccc atccccccag gggcgtgcgc ctcggccgc 11040
 tggcaggcca agcttgctt gtcgtccgc gtgaacgtcg gtcgattgtt acctgcgttc 11100
 aaatactttg cgatcggtt ggcgcctgc cgggtgcgtc ggtgatctc acggatcgac 11160
 tgcttccttc gcaacgc 11220
 ttgcccgtc gtcaccgtt ttggggggaa ggtgcacatg gtcagttct caatggaaat 11280
 tatctgcata accggctc 11340
 agcggc 11400
 tggatcagca atgagatgta tggtaat 11460
 tttcaattt aaaaatgtt atgtccgc 11520
 aaaacgacaa attacgatcc gtcgtattt taggc 11580
 ttcgaaatc ttatccgc cgtgtctaca ttcacgtcca aatggggag atccgtcgac 11640
 ctgcagttagt gtcgtccgc tgaagaagg 11700
 ccatcatcca gccagaaatg gggggagcca cgggtgatga gagcttgc 11760
 cagttgtga ttttgaactt ttgc 11820
 gtgatctgtat cttcaactc agcaaaatgtt cgatttattt aacaaagcc 11880
 caaatctct gatgttacat tgcacaagat aaaaatatacatcatgaaac 11940
 tctgc 12000
 tgctcgaggc cgcgattttt ttccaacatg gatgttgcattt tttatgggtt 12060
 cgcgataatg tcgggcaatc aggtgc 12120
 ccagagg 12180
 gtcgactaa actggctgac ggaatttgc 12240
 actcctgatg atgcattttt actcaccact gcgatccctg gggaaacacgc 12300
 ttagaagaat atcctgattt aggtaaaat attgttgcattt cgctggc 12360
 cggttgcattt cgttgcattt ttgttgcattt ctttttacatgc 12420
 gtcaggcgc aatcac 12480
 cgtatggctt ggcctgtt 12540
 cggatttgc 12600
 aaattaatag gttgttattttt gttgttgc 12660
 gccagcctat ggaactgc 12720
 aaatatggta ttgataatcc tgcatttgcattt 12780
 ttttctaat cagaattttt taatgttgc 12840
 cgggacggcg gctttgtt 12900
 atctcc 12960
 ccac 13020
 cgattcaggc 13080

aggtcgacgg atctggggga tctagcagat ccgcgagggg atcgagcccg acatatgcc 13140
 cggttcggt gcgactaaca tgagttcttg gacaaatttgc atggacacgt atgagatgat 13200
 ccaaccggag gatatacgaa agctcggtcg tgcagcaatg gaacggccaa accgtgcctt 13260
 tggccccaaag aatgagggtgc tatgcataaa ggaatctacc cgttgcgtc caacagtctc 13320
 agggtaatgtatc tctatgtatc ttaaataatg ttgtcggtat ttgtaatct catatagatt 13380
 ttcaactgtgc gacgcaaaaa tattaaataaa attattattat tatctacgtv ttgattgaga 13440
 tatcatcaat attataataaa aaatatccat taaacacgt ttgatcacaa tgacagtcaa 13500
 taatctgatt tgaatattta ttaattgtaa cgaattacat aaagatcgaa tagaaaatac 13560
 tgcactgcaa atgaaaatta acacatacta ataaatgcgt caaatatctt tgccaagatc 13620
 aagcggagtg agggcctcat atccggctc tccaccaatg ccctcgacat agatgccggg 13680
 catggctca ggcgggtt taagctcaat cccatccaa tctgaatatc ctatcccgc 13800
 cccagtcgg tctaagaacg ggtctgtcca tccacctctg ttgggggtggg cgaagaactc 13860
 cagcatgaga tccccgcgt ggaggatcat ccagccggcg tcccggaaaa cgattccgaa 13920
 gcccaacctt tcatagaagg cggcggtggc atcgaatct cgtgatggca gtttggcgt 13980
 cgcttggctg gtcatttcga acccagagt cccgctcaga agaactcgtc aagaaggcga 14040
 tagaaggcga tgcgctgca atcgggagcg gcgataccgt aaagcacgag gaagcggta 14100
 gcccattcgc cgccaagetc ttcagcaata tcacggtag ccaacgcstat gtcctgatag 14160
 cggtccgcca caccaggccg gccacagtcg atgaatccag aaaagcggcc attttccacc 14220
 atgatattcg gcaaggaggc atcggcatgg cgcgccttga gcctggcgaa cagttcggt 14280
 tcatctgtat cgacaagacc ggttccatc ggcgcgagcc cctgatgctc ttctgtccaga 14340
 gcttgggtgt cgaatggca ggtagccgg tcaagctat gcagccgccc cattgcatca 14460
 gccatgtatgg atacttttc ggcaggagca aggtgagatg acaggagatc ctgccccggc 14520
 acttcgcccc atagcagcca gtccttccc gcttcagtga caacgtcgag cacagctcg 14580
 caaggaacgc ccgtcggtgc cagccacgt agccgcgtg cctcgctctg cagttcattc 14640
 agggcaccgg acaggctcggt cttgacaaaa agaaccgggc gcccctgcgc tgacagccgg 14700
 aacacggcgg catcagagca gccgattgtc tttgtgccc agtcatagcc gaatagcctc 14760
 tccaccaag cggccggaga acctgcgtc aatccatctt gtcaatcat gctaaaggat 14820
 ctcgatcccc gggctgatcc tctcagtctc cagatgtg tttaatagg cagtagcctt 14880
 ttgatatacg ccacaagtgt gtggaaatct tatcttcgg tttcaattag gaattaacct 14940
 tattgaattc tcttggaaagg aagtccgcaa agtgggtgtc gtttcttta atgtgttcaa 15000
 catcaaatga atagtggcta agccatgtt gccatctgtat gtttcttcca agtttcaat 15060
 ctccttgtatc attgagatta acgaaactct tgaaatgagt attatctgtc ctaatcagaa 15120
 aatgaacagg agttagataa atactaaatt tctttatagt atttattacc gccaatgtct 15180
 ctttgcatt gctgtggta ttctttctg cagtttaaa gtcagatct tgcaggtccc 15240
 cagattagcc tttcaattt cagaaagaat gctaaccac agatggtagt agaggcttac 15300
 gcagcaggc tcatcaagac gatctacccg agcaataatc tcaggagat caaataccctt 15360
 cccaaagg ttaaagatgc agtcaaagaat ttcaggacta actgcataa gaacacagag 15420
 aaagatatat ttctcaagat cagaagttact attccagttt ggacgatcca aggcttgctt 15480
 cacaaccac ggcaagtaat agagattgga gtctctaaa aggttagttc cactgaatca 15540
 aaggccatgg agtcaaagat tcaaatagag gacctaaca gactcgccgt aaagactggc 15600
 gaacagttca tacagagtct cttacgactc aatgacaaga agaaaatctt cgtcaacatg 15660
 gtggagcacg acacgcttgt ctactccaaa aatatcaaag atacagtctc agaagaccaa 15720
 agggcaattt agactttca acaaaggta atatccggaa acctccctgg attccattgc 15780
 ccagctatct gtcacttat tggtaagata gtggaaaagg aagggtggctc ctacaaaatgc 15840
 catcattgcg ataaaggaaa ggcattcggtt gaagatgcct ctggccgacag tggccccaaa 15900
 gatggacccc caccacacgg gagcatcggt gaaaaagaag acgttccaaac cacgtcttca 15960
 aagcaagtgg attgatgtca tatctccact gacgtaaaggg atgacgcaca atcccactat 16020
 ctttcgcaag acccttcctc tatataagga agttcatttc atttggagag aacacggggg 16080
 actctagagg atccggatcc gtcgaccatg agcaaaaggag aagaactttt cactggagtt 16140
 gtcccaattt ttgttgaatt agatgggtat gttaatggc acaaattttc tgcgttgc 16200
 gagggtgaag gtgtatgcac atacggaaat cttaccctta aatttatttg cactactgg 16260
 aaactacctg ttccatgcc aacacttgatc actacttct ctatgggtgt tcaatgcctt 16320
 tcccgttatac cgatcatat gaagcggcac gacttctca agagcgccat gcctgaggga 16380
 tacgtgcagg agaggaccat ctcttcaag gacgacggga actacaagac gcgtgcgtaa 16440
 gtcaagttt agggagacac cctcgtaac aggtcgacgc ttaaggaaat cgatttcaag 16500
 gaggacggaa acatccctgg ccacaaggatc gaataacaact acaaactccca caacgtatac 16560
 atcacggcag acaaacaaaa gaatggaaatc aaagctaact tcaaaatttc ccacaacatt 16620
 gaagatggat ccgttcaact agcagaccat tatcaacaaa atactccaaat tggcgatggc 16680
 cctgtcctt taccagacaa ccattacgtc tcgacacaat ctgccccttc gaaagatccc 16740

aacgaaaagc gtgaccacat ggtccttctt gagtttgtaa ctgctgctgg gattacacat 16800
 ggcattggatg agctatacaa ataaagatcc tcgaatttcc ccgatcggtt aaacatttgg 16860
 caataaaagg ttcttaagatt gaatcctgtt gccgggtctt cgtatgattt catataattt 16920
 ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 16980
 tgggtttta tgatttaggt cccgcaatta tacatttaat acgcgtataga aaacaaaata 17040
 tagcgcgcaa actaggataa attatcgccc gcgggtgtcat ctatgttact agatcgatcg 17100
 ggaatttagat c 17111

<210> 32

<211> 17116

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:pAfActAf-bin

<400> 32

gatctgcaag cgatcgccggc gcgcattta aatgccccggg cgtttaaacgc cggccgcatt 60
 cccggggat ctcgatatact ttaatttcta atatatttag aaggggttta gatttatata 120
 tgacgtccat gaaaaaaaaat tttaaaacga taggccagct ccaaagaata tattttccctg 180
 gagctggtaa atattaatta gtcctctccc ttctaatata tttttcgata tttttgattc 240
 ctttttaaga ttatatagtt ccattgccaag ctgacttcct aatataatttt atgaagttag 300
 caaaaatttt aaaaagcaca cttgacttgg tatataattta taaatgttta aacttaattt 360
 cacttgtaaa actcttgcaag gatctcgata tcttaattt ctaatataatt tagaagggt 420
 tttagatttat atatgacgtc catgaaaaaa aattttaaaa cgataggcca gctccaaaga 480
 atataatttcc ctggagctgg taaatattaa tttagtcctt cccttcttaat atattttcg 540
 atattttga ttcctttta agattatata gctccatgcc aagctgactt cctaataatat 600
 tttatataagt cagcaaaatt tttaaaaagc acacttgact ttgtatataat ttataaatgt 660
 tttaaacttaa tttcacttgtt aaaactcttg caggatccac tagttcttagg ggctgcagaa 720
 ttcggcttgc tttagcactgt tcggatgttag cgtgaatagt gaatatgtt aaacattgtt 780
 atgaaagaca ctttctttca cggtctgaat taattatgtt acaattctaa tagaaaacga 900
 attaaatttac gttgaatttgtt atgaaatcta atgaaacaag ccaaccacga cgacgactaa 960
 cgttgcctgg attgactcgg tttaaatgtt ccactaaaaaa aacggagctg tcatgtaa 1020
 cgcggatcga gcaggtcaca gtcatgaagc catcaaagca aaagaactaa tccaaaggggc 1080
 tgagatgatt aatttgttta aaaatttagtt aacacgaggg aaaaggctgc ctgacagcca 1140
 ggtcacgtta tctttacctg tggtcgaaat gattcgtgtc tgtcgatttt aattttttt 1200
 ttgaaaggcc gaaaataaaag ttgtaagaga taaaccgc tatataattt catatatttt 1260
 cctctccgct ttgaaatttgc tcgttgcctt cctcaatttc atcagccgtt ttgaatctcc 1320
 ggcgacttga cagagaagaa caaggaagaa gactaagaga gaaagtaaga gataatccag 1380
 gagattcatt ctccgttttgc aatcttcctc aatctcatct tcttccgcctc ttctttcca 1440
 aggttaatagg aactttctgg atctacttttta ttgtctggat ctcgatcttgc ttttctcaat 1500
 ttcttgaga tcttggaaatttgc gtttaatttgc ggatctgtga acctccacta aatctttttgg 1560
 tttagttaga atcgatctaa gttgaccgtt cagttgactc gattatagct accagaattt 1620
 ggcttgcacct tggatggagat atccatgttca atgttacctg gaaaatgtt tgtatatgt 1680
 aaatgaaatc tgaactgtt aagtttagatt gaatctgaac actgtcaatg ttagattgaa 1740
 tctgaacact gtttaagggtt agatgaagtt tggatgttgc ttcttcgaaa ctttaggtt 1800
 tggatgttgc tacgttgaac agaaaagctat ttctgattca atcagggtttt atttgactgt 1860
 attgaactct ttgtgtgtt ttgcagctgtt agaccatgtt ccgttctgtt gaaaccccaa 1920
 cccgtgaaat caaaaaactc gacggcctgt gggcattcag tctggatcgc gaaaactgtg 1980
 gaatttgcata gctttgggtt gaaagcgttgc tacaagaaag cccggcaattt gctgtgccag 2040
 gcagttttaa cgtatcgatc gccgtatgcgc atattcgtaa ttatcgccggc aacgtctgg 2100
 atcagcgcga agtctttata ccgaaagggtt gggcaggccca gctgtatgtt ctgcgtttcg 2160
 atgcgttgcac tcattacggc aaagtgtggg tcaataatca ggaagtgtatc gagcatcagg 2220
 gcccgtatac gcccatttgc gcccgtatgtt tatttgcgggg aaaagtgtac 2280
 gtatcaccgt ttgtgttgc aacgaactgtt actggcagac tatcccgccg ggaatgggtga 2340
 ttaccgacga aaacggcaag aaaaaggctt cttacttccca tgattttttt aactatgcgg 2400
 gaatccatcg cagcgtaatc ctctacacca cggccgaaacac ctgggtggac gatattcaccg 2460
 tggtgacgca tggatgttgc aactgttgc accgttgcgtt tgactggca gttggccca 2520
 atggtgatgtt cagcgatgttgc aactgttgc accgttgcgtt tgactggca gttggccca 2580
 gcactagcgg gactttgc aactgttgc accgttgcgtt cggatcaaca gttgggttgc actggacaag 2640

tctatgaaact gtgcgtcaca gccaaaagcc agacagagtg tgatatctac ccgcttcg 2700
tcggcatccg gtcagttggca gtgaaggcg aacagttct gattaaccac aaaccgttct 2760
actttactgg ctttggtcgt catgaagatg cggaacttacg tggcaaagga ttgcataacg 2820
tgctgtatgg gcacgaccac gcattaatgg actggattgg ggccaactcc taccgtacct 2880
cgcattaccc ttacgctgaa gagatgctcg actggcaga tgaacatggc atcgtggta 2940
ttgatgaaac tgctgtgtc ggctttaacc tctcttagg cattggttc gaaggggca 3000
acaagccgaa agaactgtac agcgaagagg cagtcaacgg ggaaactcag caagcgcact 3060
tacaggcgat taaagagctg atagcgcgtg aaaaaaacc acaaagcgtg gtatgtgga 3120
gtattgccaa cgaaccggat acccgccgc aagtgcacgg gaatattcg ccactggcg 3180
aagcaacgcf taaaactcgac ccgacgcgtc cgatcacctg cgtaatgtat atgttctgcg 3240
acgctcacac cgataccatc agcgtatctt ttgatgtgtc gtgcctgaac cgatttacg 3300
gatggtatgt ccaaagcggc gattggaaa cggcagagaa ggtactggaa aaagaacttc 3360
tggcctggca ggagaaaactg catcagccga ttatcatcac cgaatacggc gtggatacgt 3420
tagccggct gcactcaatg tacaccgaca tggatgtgtc cagcggtaac aagaaaggga 3480
tggatatgtt tcaccgcgtc tttgatcgcg tattgcgcgt tggcggtaac caggtatgga 3540
atttccccga ttttgcgacc tcgcaaggca tcttcactcg cgaccgcaaa ccgaagtcgg 3600
gcatgaactt cggtaaaaaa ccgcagcagg cggctttct gctgcaaaaaa cgctgactg 3660
gcccaccatc gtcggctaca gcctcggtgg gaggcaaca atgaatcaac aacttcctg 3720
ttcaaacatt tggcaataaa gtttcttaag ggaattggag agctcgaatt tccccgatcg 3780
tatcatataa tttctgttga attacgttaa attgaatctt gttgcccgtc ttgcgtatgat 3840
gttattttat agatgggtt ttatgattag gcatgtataa attaacatgt aatgcgtac 3900
agaaaacaaa atatagcgcg caaactagga agtcccggaa ttatacattt aatacgcgt 3960
actagatcga tcgggaattt tttggggat taaattatcg cgcgcgggt catctatgtt 4020
aagggggaaa gatttataa tgacgtccat gaaaaaaat ttaaaaacga taggcagct 4080
ccaaagaata tatttcctg gagctggtaa atattaatta gtcctctccc ttctaataata 4140
ttttcgata ttttgattt cttttaaga ttatatagtc ccatgccaag ctgacttcct 4200
aatatatttt atgaagtca gaaaaattttt aaaaagcaca cttgacttgg tatataattta 4260
taaatgttta aacttaattt cacttgtaaa atattttcc ctttgcag gatctcgata tctttaattt 4320
ctaataattt tagaaggggt ttagatttt atatgtacg catgaaaaaa aattttaaaa 4380
cgataggcca gtccaaaga atatatttcc ttggagctgg taaatattaa ttatgtacg 4440
cccttctaat atattttcg atattttga aatattttt ctttgcag gatctcgata tctttaattt 4500
aagctgactt ctaataatata tttatgttca gtttgcactgttgg ttttttttta agatttataa 4560
tggtatataat ttataaatgt ttaaacttaa atattttcc ctttgcag gatctcgata tctttaattt 4620
tagttctaga gcccgcgtt aattaaggcc atattttcc ctttgcag gatctcgata tctttaattt 4680
ctacgatggg gggcatcgca ccggtgagta atattttcc ctttgcag gatctcgata tctttaattt 4740
gtagacctca attgcgagct ttcttaatttcc atattttcc ctttgcag gatctcgata tctttaattt 4800
gatgctgact ggcaggatata tttatgttca gtttgcactgttgg ttttttttta agatttataa 4860
tatgtttgtt tgattgttcc tttatgttca gtttgcactgttgg ttttttttta agatttataa 4920
ttgatttagc cctgatgaaac tgccgagggg tttatgttca gtttgcactgttgg ttttttttta agatttataa 4980
tttcgtgtca caacgagacg acagaacacc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5040
actgaaaaag gaaaattatt gtgcaggca aacccatctt gtttgcactgttgg ttttttttta agatttataa 5100
gaaatgacgg ctaaatggga aacgtattta aacccatctt gtttgcactgttgg ttttttttta agatttataa 5160
gagaacttta ttacgaatata aaaaaatttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5220
gatataaaaa aactaaattt ttctgttatttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5280
agtattttat gaaaatgtcc taagtgtggc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5340
ggttgttcaaa attatctgttca atgttaagtttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5400
ctcacaaaaaa ccaaatttca aacccatctt gtttgcactgttgg ttttttttta agatttataa 5460
aaaacgaaag atagaaagtc ctacaatgcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5520
gatttatataat ctttctcaaa ataaacataaa aacccatctt gtttgcactgttgg ttttttttta agatttataa 5580
atcacaaaaatc acttatcaca aatcacaatgttca aacccatctt gtttgcactgttgg ttttttttta agatttataa 5640
taagaagaaa tagaaagaag tgagtgttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5700
gaaatgatac accaatcagt gcaaaaaaaag aacccatctt gtttgcactgttgg ttttttttta agatttataa 5760
gtgctgactt gcaccatatc ataaaaatcg aacccatctt gtttgcactgttgg ttttttttta agatttataa 5820
aagttatggaa aataagactt agaagcaac aacccatctt gtttgcactgttgg ttttttttta agatttataa 5880
aaattttgtttaa taataggaat tgaagttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 5940
agtgattaca tgaacaaaaaa tataaaatatttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 6000
ctcaacccaaa taataaaaaca attgaatttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 6060
acaggttaaag ggcatttaac gacgaaactg aacccatctt gtttgcactgttgg ttttttttta agatttataa 6120
gaatttagaca gtcatcttcc caacttatacg aacccatctt gtttgcactgttgg ttttttttta agatttataa 6180
actttaatttcc accaagatataat tctacagtttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 6240
caattccctt accaagatataat tctacagtttcc aacccatctt gtttgcactgttgg ttttttttta agatttataa 6300

ttcgtggatgttcccttacca tttaaaggcaca caaaattatta aaaaagtggg ttttggaaagc 6360
catgcgtctg acatcttatct gattgttcaa gaaggattct acaagcgta cttggatatt 6420
caccgaacac taggggttgc cttgcacact caagtctcgaa ttccagcaatt gcttaagctg 6480
ccagcgaat gcttcatcc taaaccaaaa gtaaacagtg tcttaataaaa acttacccgc 6540
cataccacag atgttccaga taaatattgg aagctatata cgtaacttgc ttcaaaatgg 6600
gtcaatcgag aatatcgta actgtttact aaaaatcagt ttcatcaagc aatgaaacac 6660
gccaaagtaa acaatttaag taccgttact tatgagcaag tattgtctat ttttaatagt 6720
tatctattat ttaacgggag gaaataattc tatgagtcgc tttgtaaat ttggaaagtt 6780
acacgttact aaagggaatg tagataaattt attaggataa ctactgacag cttccaaggaa 6840
gctaaagagg tccctagcgc ctacgggaa tttgtatcga taagggtac aaattcccac 6900
taagcgctcg ggggctgaga aagcccagta agggaaacaac tggtaggttcg agtcgcgaga 6960
tccccccgaa cccaaaggaag taggttaaac ccgctccgat caggccgagc cacgccaggc 7020
cgagaacatt ggttccctgta ggcattcgggaa ttggcggtac aaacactaaa gctactggaa 7080
cgagcagaag tcctccggcc gccagttgccc agggcgtaaa ggtgagcaga ggcacgggag 7140
gttgccactt gcgggtcagc acgggtccga acggcatgga aaccgcccccc gccagccccg 7200
ctgcgacgccc gacaggatct agcgtctgcgt ttgggtgtcaa caccacacg cccacgccccg 7260
cagttccgca aatagcccccc aggaccgcca tcaatcgat cgggtcacct agcagagcgg 7320
cagagatgaa cacgaccatc agcgtctgca ccagaatagc gaaatattaa gtgcggcgag 7380
caggcgttag accgaaataa acaacaagct tggaatctgt cggacgatca tcacgagcaa 7440
gatgaagatg cgcatccacc agatcccgtt accggcgacc cctcggcctc gctgtcggg 7500
taaaccggcc ggcaacgccc gcagcagcat gtgcccctt gtgagcggtcc ttggggccgt ctcctgttt 7560
ctccacgaaa acgcccggaca gatgcccctt gatgtaggcg ccgaatgcca cggcatctcg 7620
gaagaccgac agccaaatgaa tctcggcgtc tttctctcg tgctcgtaaa cggacccgaa 7680
caaccgttca ggcacgcctt ccattgggtt tcggatctcg cacaattgtc aattttatc ctcgtttat 7740
catctctgga gttttcttca gggccgacaa cgctccaaacg cgtcgaatct gacgtactg agcgaagcaa 7800
cgctccaaacg cgtcgaatct gaggcttaat agcgtactg tgagcggtcc ttccagcttgc 7860
ccggcgttgc tagagcgcgc cgtgcccgg aaggcgttgc tgctgaaccc ccagccggaa 7920
cagtgcccgcc ttgttctgaa aatggcgttcc cggcggatgt ttccagcttgc agcgggtacg 8040
ctgaccggcc aaggccctag cggttgcattt cggcggatgttgc cggcggacgttgc 8100
ccaccaggcc gtcgtccgc aactcttcgc ggcggacgttgc ttccagcttgc agcgggtacg 8160
tcacgggggt ggaatccgat ccgcacatgaa tgtaataccg ggcattgtac gaccagccca 8220
gctccgggtg cgagctgaaa tagtgcgaaac cggcggatgttgc cggcggacgttgc 8280
acttctccca tatgaatttgc gtgttagtgcg cggcggatgttgc cggcggacgttgc 8340
cgatcaggac ctggcaacgg gacgtttctt ggtcggacgttgc ttccagcttgc agcgggtacg 8400
gcagcgcacac cgattccagg tgcccaacgc tgtaataccg ggcattgtac gaccagccca 8460
gttaggcgcga caggcatttc tcggccttcg tgccacgttc gtcgtccgc 8520
ggtcctggca aagctcgtag aacgtgaagg tgatcggtc ggcgataggg gtcgcgttcg 8580
cgtaactccaa cacctgctgc cacaccaggta cgtcatcgtc gggcccgacg tcgacgccccg 8640
tgttaggtgtat cttcacgtcc ttgttgcgttgc gggcggatgttgc cggcggatgttgc 8700
gcgggatttt ctgttgcgc gtggtaaca cggcggatgttgc cggcggatgttgc 8760
ctcgcatcgta gtcggccac ggcgcaatata gcttggcctc gtcgacgttgc ttggccaggt 8820
gctgtctcgta gtgtttcagc aacgcggccct tcatagttcc tcgcgtgtcg atggtcatcg 8880
cctcgccggc ggtttttcgc ttcttggtcg acttcgccaac tcgtgttcg gacgacgcga acgctccacg 8940
acttcgccaac acctgcccgc tcctgttgcg ggcggatgttgc cggcggatgttgc 9000
gcgcgggcag ggcaggggga gccagttgca cggcggatgttgc cggcggatgttgc 9060
gctggaccat cgagccgacg gacttggaaagg ttccgttgcg gcttggcctc gtcgacgttgc 9120
ttgcgtatgtt tcggcatcc tcggcgaaaa accccgctc gatcgttgc ttccctggcg 9180
ccttcggcgc aaacgtccga ttcatccacc tgatttgcg cggcggatgttgc 9240
gggcaatgtg cccttatttcc tgatttgcg cccctggcgc ttgggtgtcc agataatcca 9300
ccttatcgcc aatgaagtgc gtccgtaga ccgtctggcc gtccttctcg tacttggat 9360
tccgaatctt gcccgtcaca gataccagcg accccttgc caaataacttgc 9420
cgccctgaga gccaaaacac ttgtatgcgaa agaagtgcgttgcg gtcgttgcg 9480
catcggtcg ccactcttca ttaaccgcata tatcgaaat tgcttgcggc ttgttagaaat 9540
tgccatgacg tacctcggtg tcacgggtaa gattaccgtt aaacttggaaac tgattatggc 9600
nnctcgaaat tccctcggtc ttgccttgcg cgtcggtat gtacttcacc agctccgcga 9660
agtcgtctt cttgtatggag cgcatgggaa cgtgcttggc aatcacgcgc accccccggc 9720
cgtttagcg gctaaaaaaag tcatggctct gccctcggtc ggaccacgccc catcatgacc 9780
ttgccaagct cgtcctgtt ctcttcgtat ttgcggcagca gggcggatgttgc cgtggcatca 9840
ccgaaccgcg ccgtcgccgg gtcgtcggtg agccagagtt tcagcaggcc gcccaggccgg 9900
cccaggtcgc cattgtatgcg ggccagctcg cggacgtgtc catagtccac gacgcggcgtg 9960
atttttagc cttggccgac ggccagcagg tagggccgaca ggctcatgccc gggccggcc 9990

gccttcgtcc caatcgctct tcgttcgctc ggaaggcagt acaccttgat aggtggctg 10020
cccttcgtgg tggcttggg ttcatcagcc atccgcttgc cctcatctgt tacggccg 10080
gtagccggcc agcctcgca agcaggattc ccgttgagca ccgccagg 10140
acagtgaaga agaaacaccc gtcgcgggt gggcctactt caccatctt gcccggctg 10200
cgcgttggc tacaccaagg aaagtctaca cgaaccctt ggcaaaatcc tgtatatcg 10260
gcgaaaaagg atggatatac cggaaaaatc gctataatga ccccgaa 10320
gcggaaaaga tccgtcgacc cttccgacg ctcaccggc tgggtccct cgccgctgg 10380
ctggcgccc tctatggccc tgcaaacgcg ccagaaacgc cgtcaag 10440
caccgcggcc ggcgcggc gttgtggata ccacgcggaa aacttggccc tcactgac 10500
atgaggggcg gacgttgaca ctgggggc cgactcaccc ggcgcggcg 10560
ggggcaggct cgatttcggc cggcgcacgtg gagctggca gcctcgaaa tgacagatg 10620
cgcctgattt tacggaggt tcccacagat gatgtggaca agcttgggta taagtgc 10680
gcggtattga cacttgaggc gcgcgactac tgacagatga gggcgcgat ccttgac 10740
tgagggcag agtgatgaca gatggggc gcacctattt acatttggagg ggctgtcc 10800
aggcagaaaa tccagcattt gcaagggtt ccgcccgtt ttccggccacc gctaactgt 10860
cttttaacct gcttttaaac caatattt aaaccttgtt ttaaccagg gctgcgcct 10920
ggcgcgtgac cgccgacgcg gaaggggggt gccccccctt ctcaaccctt cccggccgc 10980
taacgcggc cttccatccc cccaggggct gccccttcg gccgcgaacg gcctcaacc 11040
aaaaatggca gccaagctt gcttggcgt tccgcgtgaa cgtcgctcg attgtac 11100
cgttcaataa ctttgcgate gtgttgcgcg cctgcgggt gcgtcggtg atctca 11160
tcgactgtt ctctcgcaac gccatccgac ggatgttgtt taaaagtccc atgtggatc 11220
ctccgtgcc coggcgctca ccgttgggg gggaaagg 11280
gaaattatct gcttaaccgg ctcagttctg ctagaaacc aacatgcaag ctccaccgg 11340
tgccaaaggcgg cagcggcggc aggatattt caattgtaaa tggttcatg tccgggaaat 11400
ctacatggat cagcaatgag tatgtggc aatatggaga aaaagaaaga gtaattacca 11460
atttttttc aattcaaaaaa ttagatgtc cgcagcgta ttataatg aaagtacatt 11520
ttgataaaac gacaattac gatccgtcgt attataaggc gaaagcaata aacaattat 11580
tctaattcgg aatattttt ttcgacgtgt ctacattcac gtccaaatgg gggagatcc 11640
tcgacctgca gtgagggtctg ctcgtgaag aagggtttgc tgactcatac caggcgtgaa 11700
tcgcccattt acccagccag aaagtgggg agccacggtt gatgagagct ttgtttagg 11760
tggaccagtt ggtgattttt aacttttgtt ttgcccacgga acggtctcg 11820
gatgcgtgat ctgatccccc aactcagcaa aagttcgatt tattcaacaa agccacgtt 11880
tgtctcaaaa tctctgtatgt tacattgcac aagataaaaa tatatcatca tgaacaataa 11940
aactgtctgc ttacataaaac agtaatacaa ggggtgttat gaggcatatt caacggaaa 12000
cgcttgcgc gaggccgcga ttaaattcca acatggatgc tgatttat ggttataat 12060
gggctcgcga taatgtcggt caatcagg 12120
atgcgcaga gttgtttctg aaacatggca aaggtagcgt tgccaatgt gttacagatg 12180
agatggtcag actaaactgg ctgacggat ttatgcctt tccgaccatc aagcatttt 12240
tccgtactcc ttagatgtgca tggttactca ccactgcgtt ccctgggaaa acagcattcc 12300
aggtatttga agaataatcc gattcagg 12360
tgcgcgggtt gcattcgatt cctgtttgtt attgtccctt taacagcgtat cgcgtatttc 12420
gtctcgctca ggcgcataa cgaatgaata acgggtttgtt tgatgcgtat gattttgatg 12480
acgagcgtaa tggctggcgtt gttgaacaag tctggaaaga aatgcataaa cttttgccc 12540
tctcaccggta ttcagtcgtc actcatggt atttctact tgataaccgtt attttgac 12600
aggggaaatt aataggttg attgtatgtt gacgagtcgg aatcgcagac cgataccagg 12660
atcttgcacg cctatggaaac tgcctcggtt agtgttctcc ttcatcag aacaggcttt 12720
ttcaaaaaata tggattgtt aatcctgata tgaataaaatt gcagtttcat ttgatgctcg 12780
atgagttttt ctaatcgaaa ttggtaatt ggtgtaca ctgcagagc attacgtga 12840
cttgacggga cggcggtttt gttgaataaa tcgaactttt gctgagttga aggtcagat 12900
cacgcattt cccgacaacg cagaccgtt cgtggcaag caaaagttca aaatcaccaa 12960
ctggtccacc tacaacaaag ctctcatcaa ccgtggctt ctcactttt ggcgttcat 13020
tggggcgatt caggcgttgc atgagtcgc aacaccttct tcacgaggca gacctcagcg 13080
cctgcaggc gacggatctg ggggatctg cagatcccg agggatcga gcccgcata 13140
tgcgggtt tgcgtcgac taacatgat tcttggacaa attgtatttgg acctgtatg 13200
atgatccaac cggaggatat agcaaagctc gttcgtcgac caatggaaacg gccaaaccgt 13260
gctttgtcc ccaagaatga ggtgtatgc atgaaggat ctacccgttgc atgtccaaaca 13320
gtctcagggt taatgtctat gtatctttaaa taatgttgta ggtatggatg aatctcatat 13380
agattttcac tgcgtcgacgc aaaaatatta aataaaatattt attattatct acgtvttgt 13440
tgagatatca tcaatattt aataaaaaata tccattaaac acgatttgat acaaataatgaca 13500
gtcaataatc tgatttgaat atttattat tgcgtcaat aactaataaa tgcaatagaa 13560
aataactgcac tgcaaatgaa aattaacaca tactaataaa atctttgcca 13620

agatcaagcg gagtgagggc ctcatatccg gtctcagttt caagcacggg atccccgaag 13680
 cgcgctccac caatgccctc gacatacatgg ccgggctcga cgctgaggac attgcctacc 13740
 tttagcatgg ttcgcgcgaa ggcttaagc tcaatcccac cccaatctga atatcctatc 13800
 ccgcgcccag tccgggttaa gaacgggtct gtccatccac ctctgttggg gtgggcaag 13860
 aactccagca tgagatcccc gcgcgtggagg atcatccagc cggcgtcccc gaaaacgatt 13920
 ccgaagccca acctttcata gaaggcggcg gtggaatcga aatctcgta tggcaggtt 13980
 ggcgtcgctt ggtcggtcat ttcaaaaccc agagtccgc tcagaagaac tcgtcaagaa 14040
 ggcgatagaa ggcgatgcgc tgcaatcg gagcggcgat accgtaaacg acgaggaagc 14100
 ggtcagccca ttgcgcgca agctttcag caatatcagc ggtagccaac gctatgtcct 14160
 gatagcggtc cgccacaccc agccggccac agtcgatgaa tccagaaaaag cggccatttt 14220
 ccaccatgat attcggcaag caggcatcgc catgggtcac gacgagatcc tcgcccgtcgg 14280
 gcatgcgcgc cttgagcctg gcgaacagtt cggctggcgc gagccctgta tgctttcgt 14340
 ccagatcatc ctgatcgaca agaccggctt ccattccgagt acgtgctcgc tcgatgcgt 14400
 gtttcgttg gtggtcgat gggcaggtag ccggatcaag cgtatgcagc cggccattt 14460
 catcagccat gatggatact ttctcgccag gagcaaggtg agatgacagg agatcctgcc 14520
 ccggcacttc gcccatacg agccagtcggc ttcccgcctt agtgacaacg tcgagcacag 14580
 ctgcgcaagg aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcgtt 14640
 cattcagggc accggacagg tcggcttga caaaaagaac cggcgcggcc tcgctgaca 14700
 gccggAACAC ggcggcatca gagcagccgaa ttgtctgttgc tgcccaatca tagccaaata 14760
 gcctctccac ccaagcgccc ggagaacctg cgtcaatcc atcttgcata atcatgctaa 14820
 aggatctcga tccccgggtt gattttctca gtctccagag atgtgtttaa ataggcagta 14880
 gcctttgat atcagccaca agtgtgtggg aatcttatct tcggatttca attagaatt 14940
 aactttattt aattctcttgc aaaggaagtc tgcttgcatttgc tgcccaatgtt 15000
 ttcaacatca aatgaataatggcata aatgcgttgc aacttgcatttgc tgcccaatgtt 15060
 cgaatctcct ttgtatttgc gattaacgaa cattttttttt atagttttta ttaccggcaa 15120
 cagaaaatgtt aacaggagttt gataaataact aaattttttt atagttttta ttaccggcaa 15180
 tgtcttttgc tatttgcgtt ggttatttttgc tgcccaatgttgc tgcccaatgtt 15240
 gtcccccaggat tagccttttc aatttcgaaatcc cttacgcaggc aggtcttcata aagacgttcc acctttccaa 15300
 gaaatgcataa cccacagatg gtttagagagg 15360
 acccgagcaaa taatctccag gagatcaaata 15420
 aaagatttcg gactaactgc atcaagaaca 15480
 gtactattcc agtatggacg attcaaggct 15540
 ttggagtttc taaaaaggta gttcccaactg 15600
 tagaggaccc aacagaactc gccgtaaaga 15660
 gactcaatgc caagaagaaa atcttcgtca 15720
 acatgggttgc gcacgcacgc cttgtctact accaaagggttgc aatttgcatttgc tgcccaatgtt 15780
 acttgccttc gcaagacccct tcctctatata aaggaagttc atttcatatttgc tgcccaatgtt 15840
 gggggactct agaggatccg gatccgtcga gagttgtccc aatttgcatttgc tgcccaatgtt 15900
 cttcaagca agtggatttgc tttcaacaaa aatgcgttgc ttttgcatttgc tgcccaatgtt 15960
 acttgccttc gcaagacccct tcctctatata aaggaagttc atttcatatttgc tgcccaatgtt 16020
 gtttgcatttgc tgcccaatgtt 16080
 ccatgagcaaa aggagaagaa cttttcactg 16140
 gagttgtccc aatttgcatttgc tgcccaatgtt 16200
 gtttgcatttgc tgcccaatgtt 16260
 gtttgcatttgc tgcccaatgtt 16320
 ggcacgactt ctcaagagc gccatgcctg 16380
 agggatacg tcaaggaggg accatcttttgc tcaaggacgc cggaaactac aagacgcgt 16440
 ctgaagtcaa gtttgcatttgc tgcccaatgtt 16500
 tcaaggaggg cggaaacatc ctcggccaca aatgcgttgc acaatctgccttgc 16560
 tatacatcac ggcagacaaa caaaagaatgc gatccgtcga ttcttgcatttgc tgcccaatgtt 16620
 acatttgcatttgc tgcccaatgtt 16680
 atggccctgt ctttttgcatttgc tgcccaatgtt 16740
 atccccacgc aagacgttgc cacatggtcc ttcttgcatttgc tgcccaatgtt 16800
 cacatggcat ggttgcatttgc tgcccaatgtt 16860
 tttggcaata aagtttgcatttgc tgcccaatgtt 16920
 aatttgcatttgc tgcccaatgtt 16980
 tgagatgggt ttttgcatttgc tgcccaatgtt 17040
 aaatatacgca gcaactacg gataaattat cggcgcgcgt gtcatttcatttgc tgcccaatgtt 17100
 gatcgaaatc tagatc tgcccaatgtt 17116

<211> 646
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 3' MAR in
pAf-Act2-Af

<400> 33
ggggatctc gatatctta atttctaata tattnagaag gggtttagat ttatatatga 60
cgtccatgaa aaaaaattt aaaacgatag gccagctca aagaatataat ttccctggag 120
ctggtaata ttaattagtc ctctcccttc taatataattt ttcgatattt ttgattcctt 180
tttaagatta tatacgctca tgccaagctg acttccta atatttatg aagttagcaa 240
aattttaaa aagcacactt gacttggtat atatttataa atgtttaaac ttaatttcac 300
ttgtaaaact cttgcaggat ctcgatatct ttaatttcta atatatttag aagggttta 360
gatttatata tgacgtccat gaaaaaaaaat tttaaaacga taggccagct ccaaagaata 420
tatccccctg gagctggtaa atattaatta gtcctctccc ttctaatata ttttcgata 480
tttttgattc ctttttaaga ttatatagtt ccatgccaag ctgacttcct aatataattt 540
atgaagttag caaaattttt aaaaagcaca cttgacttgg tatataattt taaatgttta 600
aacttaattt cacttgtaaa actcttgcag gatccactag ttctag 646

INTERNATIONAL SEARCH REPORT

I. International Application No

PCT/US 99/28123

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 7 C12N15/82 C12N15/29 C12N15/67 C12N15/63 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LIU J -W ET AL: "THE INFLUENCE OF TWO PLANT NUCLEAR MATRIX ATTACHMENT REGIONS (MARS) ON GENE EXPRESSION IN TRANSGENIC PLANTS" MOLECULAR AND GENERAL GENETICS, DE, SPRINGER VERLAG, BERLIN, vol. 39, no. 1, 1 January 1998 (1998-01-01), pages 115-123, XP002910494 ISSN: 0026-8925 the whole document --- -/- -	1-10

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the International filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the International search

Date of mailing of the International search report

29 March 2000

10/04/2000

Name and mailing address of the ISA
 European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax (+31-70) 340-3016

Authorized officer

ALCONADA RODRIG., A

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 99/28123	
---	--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BREYNE P ET AL: "Characterisation of a plant scaffold attachment region in a DNA fragment that normalises transgene expression in tobacco" PLANT CELL, US, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, no. 4, 1 April 1992 (1992-04-01), pages 463-471, XP002072399 ISSN: 1040-4651 the whole document	1-10
A	WO 97 27207 A (UNIV NORTH CAROLINA ;SPIKER STEVEN (US); ALLEN GEORGE C (US); HALL) 31 July 1997 (1997-07-31) the whole document	1-10
A	US 5 773 689 A (CHILDS LISA C ET AL) 30 June 1998 (1998-06-30) the whole document	1-10
A	RIPOLL P -J ET AL: "A new yeast artificial chromosome vector designed for gene transfer into mammalian cells" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, GB, ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 210, no. 1, 27 March 1998 (1998-03-27), pages 163-172, XP004117463 ISSN: 0378-1119 the whole document	1-10
A	VON KRIES J P ET AL: "A MATRIX-SCAFFOLD ATTACHMENT REGION BINDING PROTEIN IDENTIFICATION PURIFICATION AND MODE OF BINDING" CELL 1991, vol. 64, no. 1, 1991, pages 123-136, XP000891516 ISSN: 0092-8674 cited in the application	1-10
A	STEIN G S ET AL: "REGULATION OF TRANSCRIPTION-FACTOR ACTIVITY DURING GROWTH AND DIFFERENTIATION INVOLVEMENT OF THE NUCLEAR MATRIX IN CONCENTRATION AND LOCALIZATION OF PROMOTER BINDING PROTEINS" JOURNAL OF CELLULAR BIOCHEMISTRY 1991, vol. 47, no. 4, 1991, pages 300-305, XP000864497 ISSN: 0730-2312 cited in the application	1-10
	-/-	

INTERNATIONAL SEARCH REPORT

National Application No
PCT/US 99/28123

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SANDER M ET AL: "Drosophila topoisomerase II double-strand DNA cleavage: analysis of DNA sequence homology at the cleavage site." NUCLEIC ACIDS RESEARCH, (1985 FEB 25) 13 (4) 1057-72. , XP000891514 cited in the application</p> <hr/>	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/28123

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
WO 9727207 A	31-07-1997	US	5773695 A		30-06-1998
		AU	716202 B		24-02-2000
		AU	2246997 A		20-08-1997
		BR	9707208 A		06-04-1999
		CA	2244204 A		31-07-1997
		EP	0904276 A		31-03-1999
US 5773689 A	30-06-1998	AU	673859 B		28-11-1996
		AU	5165593 A		26-04-1994
		CA	2147006 A		14-04-1994
		EP	0663921 A		26-07-1995
		WO	9407902 A		14-04-1994